Gaussian Processes for the Analysis of Electrochemical Impedance Spectroscopy Data: Prediction, Filtering, and Active Learning

Baptiste Py[a], Adeleke Maradesa[a], Francesco Ciucci[a,b,c,d,*]

[a] Department of Mechanical and Aerospace Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China

[b] Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China

[c] HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, China

[d] HKUST Energy Institute, The Hong Kong University of Science and Technology, Hong Kong SAR, China

[*]Corresponding author: francesco.ciucci@ust.hk

Phone: +852 2358 7187

Fax: +852 2358 1543

Abstract

Electrochemical impedance spectroscopy (EIS) is a widespread characterization technique used to study electrochemical systems. However, several shortcomings still limit the application of this technique. First, EIS data is intrinsically noisy, hindering spectra regression and prediction at unknown frequencies. Second, many physicochemical properties, such as the polarization resistance, are determined through non-unique equivalent circuits. Third, probed frequencies are usually log-spaced with a fixed number of points per decade, which is not necessarily optimal. This article illustrates how Gaussian processes (GPs) can overcome these three issues by showing that GPs can successfully filter EIS data and be used to determine the polarization resistance as a stochastic variable. Lastly, a GP-based, active-learning framework is developed to select EIS frequencies optimally for quick and accurate measurements.

1   Introduction

Electrochemical impedance spectroscopy (EIS) is a widespread characterization technique that allows the study of electrochemical systems spanning the energy field (battery [1–3], supercapacitor [4–6], and solar cell [7–9]) research. EIS is also widely used in sensor development [10,11], medicine [12,13], and biology [14,15]. EIS is particularly useful because it enables the characterization of such systems to obtain important parameters like conductivities [16,17], resistances [18,19], and capacitances [20,21]. However, issues remain, as the determination of these parameters may depend closely on the equivalent circuit model (ECM) used to deconvolve the EIS spectrum [22–24]. In addition, EIS data are typically affected by experimental errors, making interpolation and prediction of the spectra at untested frequencies challenging [4,25]. Lastly, while EIS data is usually captured uniformly in the log frequency space, *i.e.*, at a fixed number of points per decade (ppd), this approach is not necessarily optimal [26–28]. This article aims at overcoming these issues using Gaussian processes (GPs).

GPs are defined as a finite or infinite set of random variables such that the joint distribution of any finite subset of these variables follows a multivariate Gaussian distribution [29]. Given an unknown function being observed, GPs can then regress and predict it using a prior distribution and a set of assumptions on the properties of the function [30]. Through GPs, the uncertainty of regression and prediction can be quantified. So far, GPs have been applied to filter data [31], predict parameters in various contexts [32], and optimize experiments in the field of active learning [33,34]. As for EIS, Liu and Ciucci leveraged GPs to deconvolve the distribution of relaxation times, an emerging EIS analysis method [35–39]. Then, Maradesa *et al*. extended this framework to constrain the DRT to be non-negative using a finite GP approximation [30]. Moreover, Ciucci used GPs complying with the Hilbert transform to evaluate EIS spectra quality [40,41].

This work develops three new ways in which GPs augment experiments. First, we show that GPs can filter and smooth noisy EIS data. Next, we detail how the polarization resistance, R_p, can be understood as a stochastic Gaussian variable, and we use this insight to predict this critical parameter. Lastly, we develop a new GP-based, active-learning acquisition method to select frequencies optimally. Our approach enables data quality improvement while bounding the experimental time. Figure 1 shows how GPs can filter EIS data and how frequencies can be selected sequentially for optimal acquisition. In short, this work lays the foundations for GP-based probabilistic methods for the analysis and prediction of EIS spectra.
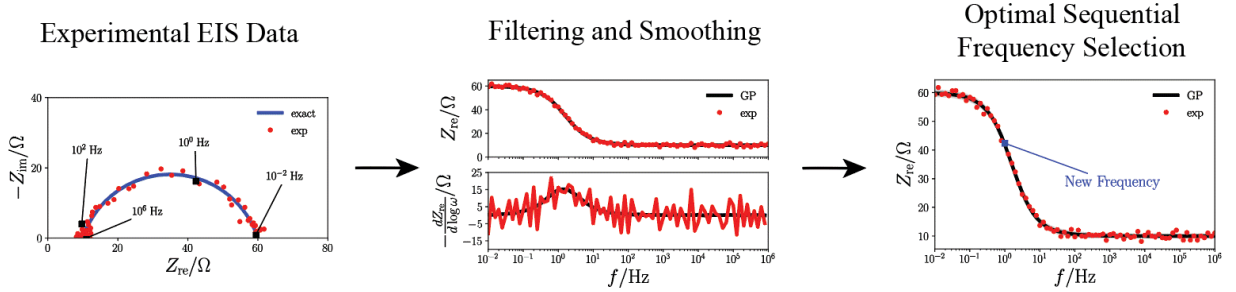

Figure 1. GPs can be used to filter an EIS spectrum and select frequencies optimally through active learning.

2 Methods

2.1 Theory

This section shows how GPs can be used to regress and predict EIS data. We also describe that, in this context, the polarization resistance is a random Gaussian variable that can be identified from data. Lastly, we detail how the GP formulation of EIS can be used to learn EIS spectra optimally.

2.1.1 Regression and Prediction with Gaussian Processes

The real, $Z_{re}(\omega)$, and imaginary, $Z_{im}(\omega)$, parts of the impedance, $Z(\omega)$, at a given angular frequency $\omega = 2\pi f$ (f is the frequency) are assumed to be joint Gaussian processes (see Section S1.1 of the supplementary information (SI) for more details) such that

$$\begin{pmatrix} Z_{re}(\omega) \\ Z_{im}(\omega) \end{pmatrix} \sim \mathcal{GP}\left(0, \begin{pmatrix} k_{re}(\omega,\omega') & k_{re-im}(\omega,\omega') \\ k_{im-re}(\omega,\omega') & k_{im}(\omega,\omega') \end{pmatrix}\right) \tag{1}$$

where the kernels $k_{re}(\omega,\omega')$, $k_{re-im}(\omega,\omega')$, $k_{im-re}(\omega,\omega')$, and $k_{im}(\omega,\omega')$ are constructed using the DRT model under the constraint of compliance with the Hilbert transform or Kramers-Kronig relations [41,42]. Hereon, kernels with band-limited (BL) and stationary-based (SB) contributions will be used [41]. For example, $k_{re}(\omega,\omega')$ is given by

$$k_{re}(\omega,\omega') = k_{re,BL}(\omega,\omega') + k_{re,SB}(\omega,\omega') \tag{2}$$

The BL kernel, $k_{re,BL}(\omega,\omega')$, is obtained from the DRT approximation and can be expressed as

$$k_{re,BL}(\omega,\omega') = \sigma_{BL}^2 \frac{\psi(\omega,\omega',\tau_{max}) - \psi(\omega,\omega',\tau_{min})}{\omega^2 - \omega'^2} \tag{3}$$

with $\psi(\omega,\omega',\tau) = \omega \arctan(\tau\omega) - \omega' \arctan(\tau\omega')$, τ_{min} and τ_{max} are the lower and upper limits of the BL kernels, respectively, and σ_{BL} is a scalar [41].

The SB kernel, $k_{re,SB}(\omega,\omega')$, is given by

$$k_{re,SB}(\omega,\omega') = 2l_{SB}^2\sigma_{SB}^2 \left(\frac{1}{2l_{SB}^2 + (\omega - \omega')^2} + \frac{1}{2l_{SB}^2 + (\omega + \omega')^2}\right) \tag{4}$$

where l_{SB} and σ_{SB} are two scalars. More details on the kernels used are provided in Section S1.2.

Both real and imaginary parts of the impedance spectrum are used for regression and prediction purposes, *i.e.*,

$$
\begin{pmatrix} Z_{\mathrm{re}}(\omega_*) \\ Z_{\mathrm{im}}(\omega_*) \\ \boldsymbol{Z}_{\mathrm{re}}(\boldsymbol{\omega}) \\ \boldsymbol{Z}_{\mathrm{im}}(\boldsymbol{\omega}) \end{pmatrix} \sim \mathcal{N}\left(0, \begin{pmatrix} \mathbf{K}(\omega_*, \omega_*) & \mathbf{K}(\omega_*, \boldsymbol{\omega}) \\ \mathbf{K}^{\mathsf{T}}(\omega_*, \boldsymbol{\omega}) & \mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega}) \end{pmatrix} \right) \tag{5}
$$

where ω_* is an angular frequency at which we wish to jointly predict or regress the real and imaginary parts of the impedance, $\boldsymbol{Z}_{\mathrm{re}}(\boldsymbol{\omega})$ and $\boldsymbol{Z}_{\mathrm{im}}(\boldsymbol{\omega})$ are the vectors of the real and imaginary parts of the experimental impedance, respectively, and $\boldsymbol{\omega}$ is the set of probed angular frequencies. The components of the covariance matrix in (5) are given by

$$
\mathbf{K}(\omega_*, \omega_*) = \begin{pmatrix} k_{\mathrm{re}}(\omega_*, \omega_*) + \sigma_{\mathrm{R}}^2 & k_{\mathrm{re-im}}(\omega_*, \omega_*) \\ k_{\mathrm{im-re}}(\omega_*, \omega_*) & k_{\mathrm{im}}(\omega_*, \omega_*) + \sigma_{\mathrm{L}}^2 \omega_*^2 \end{pmatrix} \tag{6a}
$$

$$
\mathbf{K}(\omega_*, \boldsymbol{\omega}) = \begin{pmatrix} \mathbf{k}_{\mathrm{re}}(\omega_*, \boldsymbol{\omega}) + \sigma_{\mathrm{R}}^2 \mathbf{1} & \mathbf{k}_{\mathrm{re-im}}(\omega_*, \boldsymbol{\omega}) \\ \mathbf{k}_{\mathrm{im-re}}(\omega_*, \boldsymbol{\omega}) & \mathbf{k}_{\mathrm{im}}(\omega_*, \boldsymbol{\omega}) + \sigma_{\mathrm{L}}^2 \omega_* \boldsymbol{\omega} \end{pmatrix} \tag{6b}
$$

$$
\mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega}) = \begin{pmatrix} \mathbf{k}_{\mathrm{re}}(\boldsymbol{\omega}, \boldsymbol{\omega}) + \sigma_{\mathrm{R}}^2 \mathbf{1}^{\mathsf{T}} \mathbf{1} + \sigma_n^2 \boldsymbol{I} & \mathbf{k}_{\mathrm{re-im}}(\boldsymbol{\omega}, \boldsymbol{\omega}) \\ \mathbf{k}_{\mathrm{re-im}}^{\mathsf{T}}(\boldsymbol{\omega}, \boldsymbol{\omega}) & \mathbf{k}_{\mathrm{im}}(\boldsymbol{\omega}, \boldsymbol{\omega}) + \sigma_{\mathrm{L}}^2 \boldsymbol{\omega}\boldsymbol{\omega}^{\mathsf{T}} + \sigma_n^2 \boldsymbol{I} \end{pmatrix} \tag{6c}
$$

where $\boldsymbol{I}$ is the $M \times M$ identity matrix (M is the number of frequencies), $\mathbf{1}$ is the $1 \times M$ vector whose components are all equal to 1, σ_{R} and σ_{L} are the standard deviations of the resistance and inductance, respectively [35], and σ_n is the standard deviation of the random error, $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \sigma_n^2 \boldsymbol{I})$, that is assumed to corrupt the exact impedance, $\boldsymbol{Z}_{\mathrm{exact}}$, to obtain the experimental impedance, $\boldsymbol{Z}_{\mathrm{exp}}$ [30]. Formally, we set

$$
\boldsymbol{Z}_{\mathrm{exp}} = \boldsymbol{Z}_{\mathrm{exact}} + \boldsymbol{\varepsilon} \tag{7}
$$

As for the vectors $\mathbf{k}_{\mathrm{re}}(\omega_*, \boldsymbol{\omega})$, $\mathbf{k}_{\mathrm{re-im}}(\omega_*, \boldsymbol{\omega})$, $\mathbf{k}_{\mathrm{im-re}}(\omega_*, \boldsymbol{\omega})$, $\mathbf{k}_{\mathrm{im}}(\omega_*, \boldsymbol{\omega})$, and the matrices $\mathbf{k}_{\mathrm{re}}(\boldsymbol{\omega}, \boldsymbol{\omega})$, $\mathbf{k}_{\mathrm{re-im}}(\boldsymbol{\omega}, \boldsymbol{\omega})$, $\mathbf{k}_{\mathrm{im}}(\boldsymbol{\omega}, \boldsymbol{\omega})$, their expressions are given in Table S1.

We now define the predicted impedance to be $Z_* = \begin{pmatrix} Z_{\text{re}}(\omega_*) \\ Z_{\text{im}}(\omega_*) \end{pmatrix}$ and the experimental data $Z = \begin{pmatrix} Z_{\text{re}}(\omega) \\ Z_{\text{im}}(\omega) \end{pmatrix}$. It follows from (5), that $Z_*|Z = Z_{\text{exp}}$ (*i.e.*, Z_* conditioned to $Z = Z_{\text{exp}}$) is the following multivariate normal distribution [29]

$$Z_*|Z = Z_{\text{exp}} \sim \mathcal{N}(\boldsymbol{\mu}(\omega_*), \boldsymbol{\Sigma}(\omega_*)) \tag{8}$$

where[1]

$$\boldsymbol{\mu}(\omega_*) = \mathbf{K}(\omega_*, \boldsymbol{\omega})\mathbf{K}^{-1}(\boldsymbol{\omega}, \boldsymbol{\omega})Z \tag{9a}$$

$$\boldsymbol{\Sigma}(\omega_*) = \mathbf{K}(\omega_*, \omega_*) - \mathbf{K}(\omega_*, \boldsymbol{\omega})\mathbf{K}^{-1}(\boldsymbol{\omega}, \boldsymbol{\omega})\mathbf{K}^{\top}(\omega_*, \boldsymbol{\omega}) \tag{9b}$$

Leveraging (9a), the out-of-sample mean square error, OMSE, between the exact impedance and the impedance regressed with GPs, can be defined as

$$\text{OMSE} = \int_{-\infty}^{+\infty} |Z_{\text{exact}}(\omega) - \mu_Z(\omega)|^2 \text{dlog } \omega \tag{10}$$

where $\mu_Z(\omega) = \mu_{\text{re}}(\omega) + i\mu_{\text{im}}(\omega)$ with $\mu_{\text{re}}(\omega)$ and $\mu_{\text{im}}(\omega)$ the real and imaginary parts of the mean according to (9a), respectively. Conceptually, the OMSE evaluates how well the GP mean matches the exact impedance. The out-of-sample mean square derivative error, OMSE_∂, is given by

$$\text{OMSE}_\partial = \int_{-\infty}^{+\infty} \left| \frac{\text{d}Z_{\text{exact}}(\omega)}{\text{dlog } \omega} - \frac{\text{d}\mu_Z(\omega)}{\text{dlog } \omega} \right|^2 \text{dlog } \omega \tag{11}$$

---

[1] With an abuse of notation, $\mathbf{K}^{-1}(\boldsymbol{\omega}, \boldsymbol{\omega})$ and $\mathbf{K}^{\top}(\omega_*, \boldsymbol{\omega})$ refer to the inverse of the matrix $\mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega})$ and the transpose of the matrix $\mathbf{K}(\omega_*, \boldsymbol{\omega})$, respectively.

The experimental OMSE, OMSE_{exp}, and the experimental OMSE_{∂}, $\text{OMSE}_{\partial,\text{exp}}$, can be loosely defined as

$$\text{OMSE}_{\text{exp}} = \mathbb{E}\left(\int_{-\infty}^{+\infty} \left| Z_{\text{exact}}(\omega) - Z_{\text{exp}}(\omega) \right|^2 \text{dlog}\,\omega \right) \qquad (12a)$$

$$\text{OMSE}_{\partial,\text{exp}} = \mathbb{E}\left(\int_{-\infty}^{+\infty} \left| \frac{\text{d}Z_{\text{exact}}(\omega)}{\text{dlog}\,\omega} - \frac{\text{d}Z_{\text{exp}}(\omega)}{\text{dlog}\,\omega} \right|^2 \text{dlog}\,\omega \right) \qquad (12b)$$

with $\mathbb{E}[\cdot]$ the expectation. From (7), it follows that OMSE_{exp} and $\text{OMSE}_{\partial,\text{exp}}$ are the expectations of the random variable $|\varepsilon|^2$ and that of the squared of the first derivative of ε, respectively. Therefore, both metrics depend only on the error model, and the integrals in (12) diverge. They converge, however, to two finite values if discretized. More details are given in Section 2.2.1. To assess filtering performance, OMSE and OMSE_{exp} were compared, while we utilized OMSE_{∂} and $\text{OMSE}_{\partial,\text{exp}}$ to evaluate smoothing (see Section 3.1.1).

2.1.2 Evaluation of the Polarization Resistance

The polarization resistance, R_{p}, is given by [43] [2]

$$R_{\text{p}} = Z_{\text{re}}(0) - Z_{\text{re}}(\infty) \qquad (13)$$

where R_{p} corresponds loosely to the diameter of the EIS spectrum in the Nyquist plot [44].

Thus, we aim to estimate $Z_{\text{re}}(0)$ and $Z_{\text{re}}(\infty)$, where

$$\begin{pmatrix} Z_{\text{re}}(0) \\ Z_{\text{re}}(\infty) \\ \mathbf{Z}_{\text{re}}(\boldsymbol{\omega}) \\ \mathbf{Z}_{\text{im}}(\boldsymbol{\omega}) \end{pmatrix} \sim \mathcal{N}\left(0, \begin{pmatrix} \mathbf{K}_{\text{re}}(0,\infty) & \mathbf{K}_{\text{re}-\text{im}}(0,\infty) \\ \mathbf{K}_{\text{re}-\text{im}}^{\text{T}}(0,\infty) & \mathbf{K}(\boldsymbol{\omega},\boldsymbol{\omega}) \end{pmatrix} \right) \qquad (14)$$

---

[2] With an abuse of notation, ∞ is to be understood as the limit for $\omega_* \to \infty$.

and

$$\mathbf{K}_{\mathrm{re}}(0,\infty) = \begin{pmatrix} k_{\mathrm{re}}(0,0) + \sigma_{\mathrm{R}}^2 & k_{\mathrm{re}}(0,\infty) + \sigma_{\mathrm{R}}^2 \\ k_{\mathrm{re}}(\infty,0) + \sigma_{\mathrm{R}}^2 & k_{\mathrm{re}}(\infty,\infty) + \sigma_{\mathrm{R}}^2 \end{pmatrix} \tag{15a}$$

$$\mathbf{K}_{\mathrm{re-im}}(0,\infty) = \begin{pmatrix} \mathbf{k}_{\mathrm{re}}(0,\boldsymbol{\omega}) + \sigma_{\mathrm{R}}^2\mathbf{1} & \mathbf{k}_{\mathrm{re-im}}(0,\boldsymbol{\omega}) \\ \mathbf{k}_{\mathrm{re}}(\infty,\boldsymbol{\omega}) + \sigma_{\mathrm{R}}^2\mathbf{1} & \mathbf{k}_{\mathrm{re-im}}(\infty,\boldsymbol{\omega}) \end{pmatrix} \tag{15b}$$

Table 1 presents the limits of each kernel used in the covariance matrices in (15).

Table 1 - Limits of the kernels $k_{\mathrm{re}}(\omega,\omega')$ and $k_{\mathrm{re-im}}(\omega,\omega')$ used in (15).

ω	ω'	$k_{\mathrm{re}}(\omega,\omega')$	$k_{\mathrm{re-im}}(\omega,\omega')$
0	0	$(\tau_{\max} - \tau_{\min})\sigma_{\mathrm{BL}}^2 + 2\sigma_{\mathrm{SB}}^2$	0
0	finite and non-zero	$\arctan\left(\dfrac{\tau_{\max} - \tau_{\min}}{1 + \tau_{\max}\tau_{\min}\omega'^2}\omega'\right)\dfrac{\sigma_{\mathrm{BL}}^2}{\omega'}$ $+ \dfrac{4l_{\mathrm{SB}}^2\sigma_{\mathrm{SB}}^2}{2l_{\mathrm{SB}}^2 + \omega'^2}$	$-\ln\left(\dfrac{1 + \omega'^2\tau_{\max}^2}{1 + \omega'^2\tau_{\min}^2}\right)\dfrac{\sigma_{\mathrm{BL}}^2}{2\omega'}$ $+ \dfrac{4\sqrt{2}l_{\mathrm{SB}}\sigma_{\mathrm{SB}}^2\omega'}{2l_{\mathrm{SB}}^2 + \omega'^2}$
0	∞	0	0
∞	finite	0	0
∞	∞	σ_{SB}^2	0

Based on the same procedure used in Section 2.1.1, it follows that $\begin{pmatrix} Z_{\mathrm{re}}(0) \\ Z_{\mathrm{re}}(\infty) \end{pmatrix}|Z = Z_{\mathrm{exp}}$ is the

following multivariate normal distribution:

$$\begin{pmatrix} Z_{\mathrm{re}}(0) \\ Z_{\mathrm{re}}(\infty) \end{pmatrix}|Z = Z_{\mathrm{exp}} \sim \mathcal{N}\big(\boldsymbol{\mu}_{\mathrm{re}}(0,\infty), \boldsymbol{\Sigma}_{\mathrm{re}}(0,\infty)\big) \tag{16}$$

where the mean, $\boldsymbol{\mu}_{\mathrm{re}}(0,\infty) = \begin{pmatrix} \mu_{\mathrm{re}}^0 \\ \mu_{\mathrm{re}}^\infty \end{pmatrix}$, and covariance, $\boldsymbol{\Sigma}_{\mathrm{re}}(0,\infty) = \begin{pmatrix} \Sigma_{\mathrm{re}}^0 & \Sigma_{\mathrm{re}}^{0,\infty} \\ \Sigma_{\mathrm{re}}^{0,\infty} & \Sigma_{\mathrm{re}}^\infty \end{pmatrix}$, are defined as

$$\boldsymbol{\mu}_{\mathrm{re}}(0,\infty) = \mathbf{K}_{\mathrm{re-im}}(0,\infty)\mathbf{K}^{-1}(\boldsymbol{\omega},\boldsymbol{\omega})Z \tag{17a}$$

$$\boldsymbol{\Sigma}_{\mathrm{re}}(0,\infty) = \mathbf{K}_{\mathrm{re}}(0,\infty) - \mathbf{K}_{\mathrm{re-im}}(0,\infty)\mathbf{K}^{-1}(\boldsymbol{\omega},\boldsymbol{\omega})\mathbf{K}_{\mathrm{re-im}}^{\mathsf{T}}(0,\infty) \tag{17b}$$

Given (13) and the derivation provided in Section S1.3, R_{p} is a Gaussian random variable

$$R_{\mathrm{p}} \sim \mathcal{N}\left(\mu_{R_{\mathrm{p}}}, \sigma_{R_{\mathrm{p}}}^{2}\right) \tag{18}$$

with $\mu_{R_{\mathrm{p}}} = \mu_{\mathrm{re}}^{0} - \mu_{\mathrm{re}}^{\infty}$ and $\sigma_{R_{\mathrm{p}}} = \sqrt{\Sigma_{\mathrm{re}}^{0} - 2\Sigma_{\mathrm{re}}^{0,\infty} + \Sigma_{\mathrm{re}}^{\infty}}$.

2.1.3   Active Learning with Gaussian Processes

Active learning is the field of machine learning dedicated to identifying the input parameters one at a time to augment computer simulations [45]. Two standard methods are the active learning MacKay (ALM) [34] and active learning Cohn (ALC) [46]. While the ALM aims at maximizing the predicted variance, the ALC targets the maximization of the expected average variance. The next two sections define the ALM and ALC in the context of EIS.

2.1.3.1   Active Learning MacKay

The ALM design selects a new angular frequency, ω_{new}, given a set of previously probed angular frequencies, $\boldsymbol{\omega}_{\mathrm{old}}$, as follows

$$\omega_{\mathrm{new}} = \underset{\omega}{\mathrm{argmax}}\, \phi\big(\boldsymbol{\Sigma}_{\mathrm{ALM}}(\omega|\boldsymbol{\omega}_{\mathrm{old}})\big) \tag{19}$$

where the covariance matrix, $\boldsymbol{\Sigma}_{\mathrm{ALM}}$, is given as in (9b) by the following expression:

$$\boldsymbol{\Sigma}_{\mathrm{ALM}}(\omega|\boldsymbol{\omega}_{\mathrm{old}}) = \mathbf{K}(\omega,\omega) - \mathbf{K}(\omega,\boldsymbol{\omega}_{\mathrm{old}})\mathbf{K}^{-1}(\boldsymbol{\omega}_{\mathrm{old}},\boldsymbol{\omega}_{\mathrm{old}})\mathbf{K}^{\mathsf{T}}(\omega,\boldsymbol{\omega}_{\mathrm{old}}) \tag{20}$$

and $\phi(\boldsymbol{\Sigma}_{\mathrm{ALM}})$ is either the trace (A-criterion), the determinant (D-criterion), or the maximum of the eigenvalues (E-criterion) of $\boldsymbol{\Sigma}_{\mathrm{ALM}}$ [47].

2.1.3.2   Active Learning Cohn

In the ALC framework, the subsequent angular frequency, ω_{new}, is obtained as

$$\omega_{\mathrm{new}} = \underset{\omega}{\mathrm{argmin}} \int_{-\infty}^{+\infty} \phi\big(\boldsymbol{\Sigma}_{\mathrm{ALC}}(\omega'|\boldsymbol{\omega})\big)\mathrm{dlog}\,\omega' \tag{21}$$

where $\boldsymbol{\omega} = (\omega, \boldsymbol{\omega}_{\mathrm{old}})$, ω' is an angular frequency within the integration range, and $\boldsymbol{\Sigma}_{\mathrm{ALC}}$ is the ALC covariance matrix given by

$$\boldsymbol{\Sigma}_{\mathrm{ALC}}(\omega'|\boldsymbol{\omega}) = \mathbf{K}(\omega', \omega') - \mathbf{K}(\omega', \boldsymbol{\omega})\mathbf{K}^{-1}(\boldsymbol{\omega}, \boldsymbol{\omega})\mathbf{K}^{\top}(\omega', \boldsymbol{\omega}) \tag{22}$$

More details on the numerical integration of (21) are given in Section 2.2.2.

2.2 Implementation

2.2.1 Synthetic EIS experiments

Synthetic experiments were used to assess the consistency of the methods developed in this article. To replicate the conditions of real EIS experiments, and unless otherwise specified, the log-spaced frequencies for the synthetic EIS experiments ranged from 10^{-2} to 10^{6} Hz with 10 ppd, *i.e.*, 81 frequencies [26,48]. Different models for the exact impedance, $Z_{\mathrm{exact}}(\omega)$, were considered, namely the 1×ZARC, 2×ZARC, Gerischer, hook, Havriliak-Negami, fractal, and piecewise constant (PWC) models [30,49,50]. Analytical expressions for the impedance of each model are reported in Table 2. The parameters used are given in Table S2.

The GP hyperparameters (*i.e.*, σ_{BL}, σ_{SB}, l_{SB}, σ_{R}, σ_{L}, and σ_{n}) were obtained by maximizing the experimental evidence, as illustrated in [30] and [35].

When needed, the covariance matrix, $\mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega})$, was replaced by its nearest positive definite matrix so that the inverse, $\mathbf{K}^{-1}(\boldsymbol{\omega}, \boldsymbol{\omega})$, could be computed using Cholesky decomposition [30,51].

OMSE, $\mathrm{OMSE}_{\mathrm{exp}}$, OMSE_{∂}, and $\mathrm{OMSE}_{\partial,\mathrm{exp}}$ were computed numerically (unless otherwise specified) using the trapezoidal rule from $2\pi \times 10^{-2}$ to $2\pi \times 10^{6}$ rad/s with 20 ppd.[3] Although GPs can be

---

[3] We must remind the reader that the frequency, f (Hz), and the angular frequency, $\omega = 2\pi f$ (rad/s), are different quantities.

implemented to regress derivatives, *e.g.*, in [52], we simply used the numerical derivatives of Z_{exact}, Z_{exp}, $\mu_{\text{re}}(\omega)$, and $\mu_{\text{im}}(\omega)$ with respect to $\log \omega$ in order to estimate OMSE_{∂} and $\text{OMSE}_{\partial,\text{exp}}$.

Moreover, 100 synthetic EIS spectra were generated for each case. For all boxplots, the notch refers to the median, the box to the interquartile range, the circles to the outliers, and the whiskers to the interquartile range multiplied by 1.5. In the Bode plots, the means and $3 \times \sigma$ credible bands for the real and imaginary parts of the impedance are shown as black lines and gray regions, respectively.

Table 2 – Exact impedance for each impedance model studied in this work where R_∞, R_{ct}, $R_{\infty,1}$, $R_{\text{ct},1}$, $R_{\infty,2}$, and $R_{\text{ct},2}$ are resistances; τ_0, τ_1, and τ_2 are timescales; ϕ, ϕ_1, and ϕ_2 are dispersion factors; and L_0 is an inductance.

Model	Exact impedance $Z_{\text{exact}}(\omega)$	Reference
1×ZARC with an inductance	$R_\infty + \dfrac{R_{\text{ct}}}{1 + (i\omega\tau_0)^\phi} + i\omega L_0$	[35]
2×ZARC and hook	$R_{\infty,1} + \dfrac{R_{\text{ct},1}}{1 + (i\omega\tau_1)^{\phi_1}} + R_{\infty,2} + \dfrac{R_{\text{ct},2}}{1 + (i\omega\tau_2)^{\phi_2}}$	[30,53]
Gerischer	$R_\infty + \dfrac{R_{\text{ct}}}{\sqrt{1 + i\omega\tau_0}}$	[50]
Havriliak-Negami	$R_\infty + \dfrac{R_{\text{ct}}}{(1 + (i\omega\tau_0)^\phi)^\psi}$	[49]
Fractal	$R_\infty + \dfrac{R_{\text{ct}}}{(1 + i\omega\tau_0)^\phi}$	[54]
PWC	$R_\infty + \dfrac{R_{\text{ct}}}{\log\left(\frac{\tau_1}{\tau_0}\right)}\left(\log\left(1 - \dfrac{i}{\omega\tau_0}\right) - \log\left(1 - \dfrac{i}{\omega\tau_1}\right)\right)$	[55]

2.2.2 Active Learning

Hereon, $\boldsymbol{\omega}_{5\text{ppd}}$ refers to the log-spaced range from $2\pi \times 10^{-2}$ to $2\pi \times 10^6$ rad/s with 5 ppd ($M = 41$ frequencies). To determine an optimal frequency acquisition strategy, we implemented active

learning with the ALM and ALC for $\sigma_n^{\text{exp}} = 0.3\ \Omega$ and the 1×ZARC model with an inductance. Based on the discussion in Section 2.1.3, we started by setting $\boldsymbol{\omega}_{\text{old}} = \boldsymbol{\omega}_{\text{5ppd}}$. Then, we added at each iteration (100 iterations in total) a new angular frequency ω_{new} to $\boldsymbol{\omega}_{\text{old}}$. The new angular frequency ω_{new} was obtained by maximizing (19) for the ALM or by minimizing (21) for the ALC. The upper bound angular frequency for the ALM and ALC optimizations was set to $2\pi \times 10^6$ rad/s. We also constrained the experimental time such that $\frac{2\pi}{\omega_{\text{new}}} +$

$2\pi \sum_{k=1}^{\dim(\boldsymbol{\omega}_{old})} \frac{1}{(\boldsymbol{\omega}_{\text{old}})_k} \leq \alpha t_{\text{exp,10ppd}}$ where $\alpha = 60\%, 70\%, \ldots,$ or 100%, and $t_{\text{exp,10ppd}}$ is the time for the 10 ppd experiment. For the ALC, the integral in (21) was computed with the trapezoidal rule on the log-spaced range from $2\pi \times 10^{-2}$ to $2\pi \times 10^6$ rad/s with 10 ppd. Moreover, we updated the vector of hyperparameters only at the 30th iteration.

To compare the ALM and ALC frequency placements, we computed at each iteration the OMSE, see (10), and σ_{avg}^2 and σ_{max}^2, which are defined as

$$\sigma_{\text{avg}}^2 = \int_{-\infty}^{+\infty} \text{tr}\big(\boldsymbol{\Sigma}(\omega_*)\big) \text{dlog}\, \omega_* \tag{23a}$$

$$\sigma_{\text{max}}^2 = \max_{\omega_* \in \boldsymbol{\omega}_{\text{40ppd}}} \text{tr}\big(\boldsymbol{\Sigma}(\omega_*)\big) \tag{23b}$$

where $\boldsymbol{\Sigma}(\omega_*)$ is defined in (9b) with $\boldsymbol{\omega}$ being the current set of angular frequencies, $i.e.$, $\boldsymbol{\omega} = (\omega_{\text{new}}, \boldsymbol{\omega}_{\text{old}})$, and $\boldsymbol{\omega}_{\text{40ppd}}$ is the log-spaced range from $2\pi \times 10^{-2}$ to $2\pi \times 10^6$ rad/s with 40 ppd. The integral in (23a) was approximated using the trapezoidal rule on $\boldsymbol{\omega}_{\text{40ppd}}$. Conceptually, σ_{avg}^2 and σ_{max}^2 evaluate the GP regression average and maximum errors, respectively.

3 Results

This section illustrates how we use GPs to: 1) filter and smooth noisy EIS data; 2) compute R_{p}

and its uncertainty using an array of synthetic and real EIS experiments; and 3) select frequencies through active learning to ensure low variances and a short experimental time.

3.1 Filtering and Smoothing Noisy EIS Data with Gaussian Processes

3.1.1 Filtering and Smoothing

To assess the ability of GPs to filter noisy EIS data, we first drew 100 artificial experiments with each impedance model in Table 2. For illustration purposes, panels (a), (d), and (g) of Figure 2 show the Nyquist plot of the exact, GP-regressed, and experimental impedances for the 1×ZARC, separated 2×ZARC ($R_{ct,1} = R_{ct,2}$), and distant 2×ZARC ($R_{ct,1} < R_{ct,2}$) models, respectively. Panels (b), (e), and (h) of Figure 2 present the respective Bode plots, while the Bode plots for the hook, Gerischer, Havriliak-Negami, fractal, and PWC models are shown in Figures S1, S2, S3, S4, and S5, respectively. In all cases, the $3 \times \sigma$ credible bands for the real and imaginary parts of the impedance are shown. The OMSE values are shown in aggregate form as boxplots in panels (a) and (b) of Figure 3 for $\sigma_n^{\exp} = 0.1, 0.2,\dots, 1.0\ \Omega$ (100 synthetic experiments for each $\sigma_n^{\exp}$) for the 1×ZARC and separated 2×ZARC models, respectively. The OMSE for each model and $\sigma_n^{\exp} = 0.1, 0.3, 0.5,$ and $1.0\ \Omega$ are shown in panels (a), (b), (c), and (d) of Figure 4, respectively. Given $\sigma_n^{\exp}$, similar values of OMSE were obtained for all models. Consistent with intuition, OMSE increased monotonically with $\sigma_n^{\exp}$ for all models. However, OMSE were smaller than the experimental $OMSE_{exp}$ (see the definition in (12a)) for a given level of noise of $\sigma_n^{\exp}$ (see the boxplots of $OMSE_{exp}$ in Figures S6 and S7).
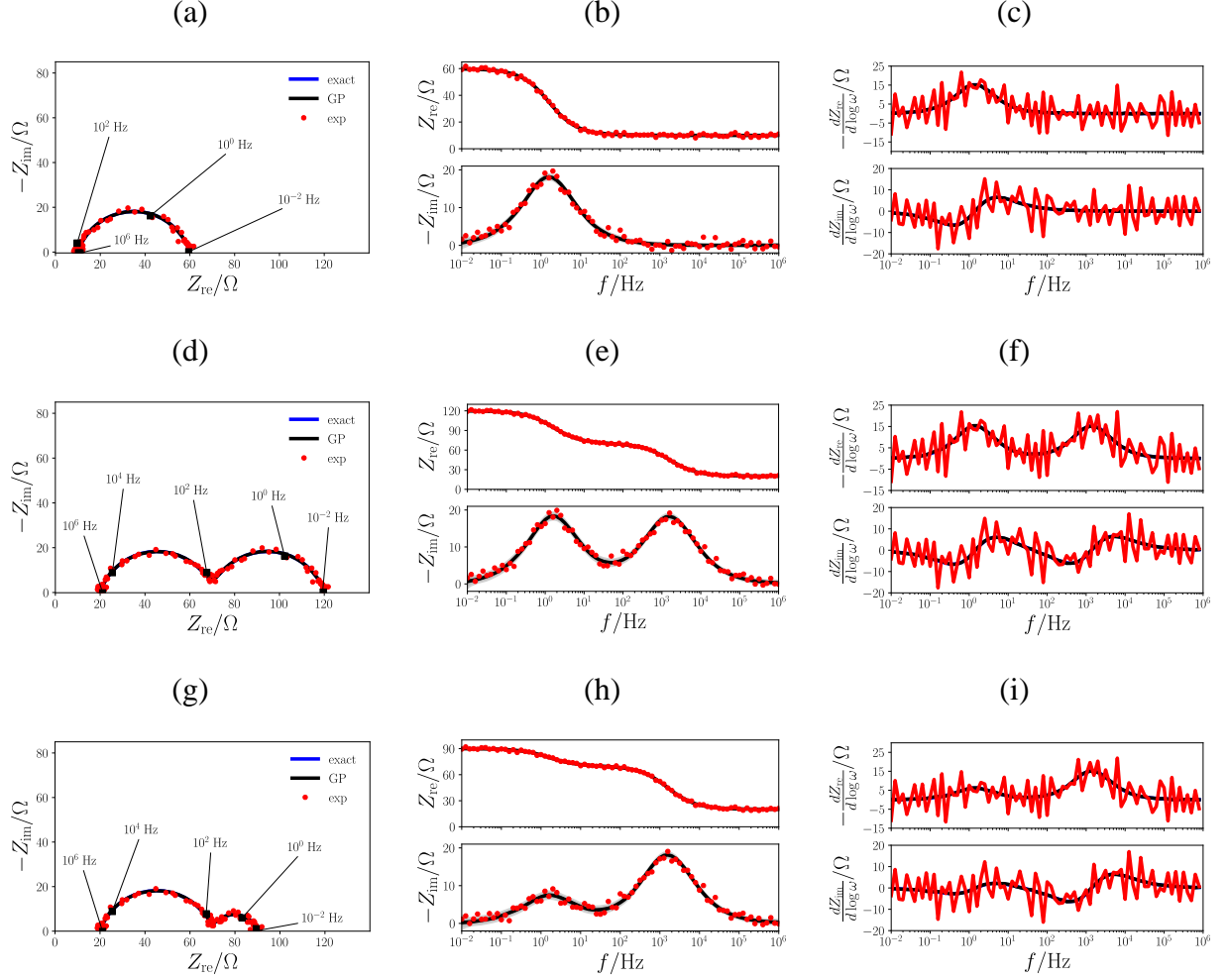
Figure 2. Nyquist plot of the exact, GP-regressed, and experimental impedances for the (a) 1×ZARC, (d) separated 2×ZARC, and (g) distant 2×ZARC models with $\sigma_n^{\text{exp}} = 1.0\,\Omega$. The corresponding Bode plots are shown in panels (b), (e), and (h). The corresponding first derivatives are shown in panels (c), (f), and (i).
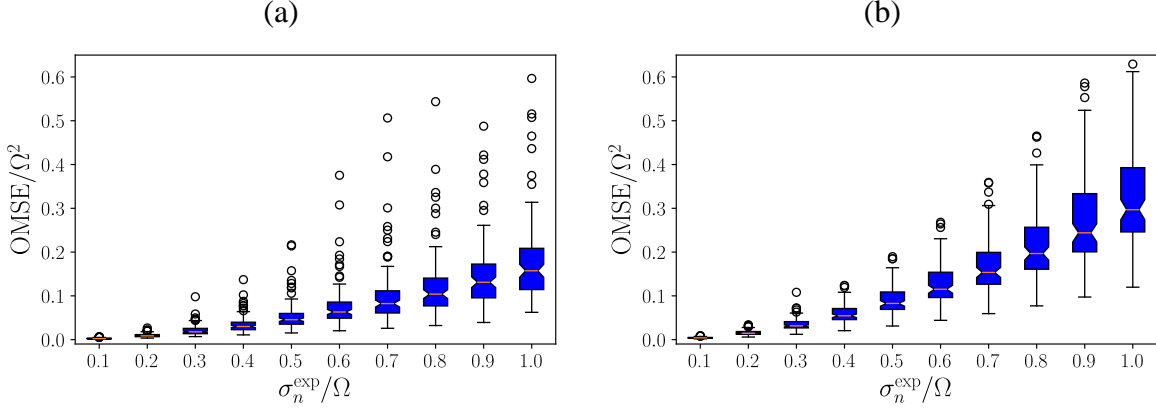
Figure 3. Boxplots of OMSE as functions of σ_n^{exp} for the (a) 1×ZARC and (b) separated 2×ZARC models.
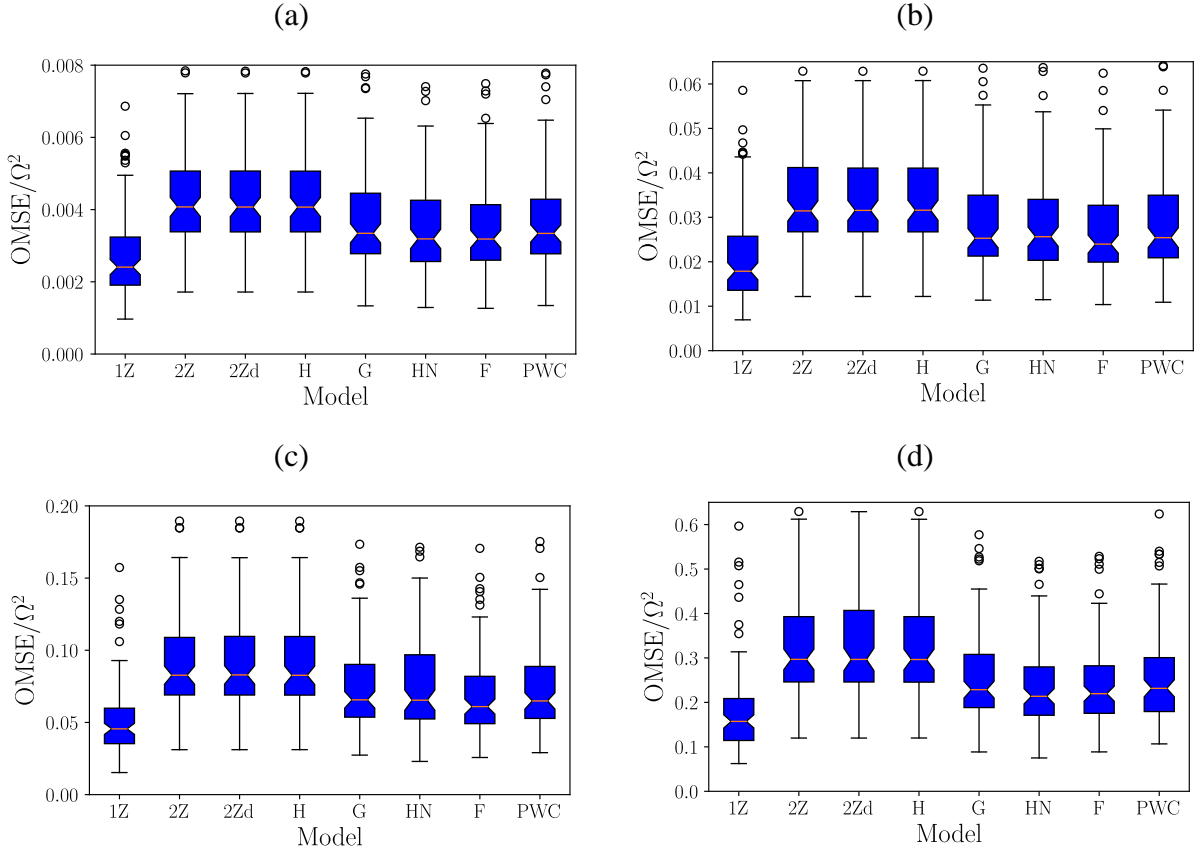
Figure 4. Boxplots of OMSE for the 1×ZARC (1Z), separated 2×ZARC (2Z), distant 2×ZARC (2Zd), hook (H), Gerischer (G), Havriliak-Negami (HN), fractal (F), and PWC models for $\sigma_n^{\text{exp}} =$ (a) 0.1, (b) 0.3, (c) 0.5, and (d) 1.0 Ω.

16

Next, we focused on GP smoothing. Given $\sigma_n^{\text{exp}} = 1.0\ \Omega$, we plotted the first derivatives of the real and imaginary parts of the exact, GP-regressed, and experimental impedances for the 1×ZARC, separated 2×ZARC, and distant 2×ZARC models in panels (c), (f), and (i) of Figure 2, respectively. The first derivatives of the real and imaginary parts of the exact, GP-regressed, and experimental impedances for the hook, Gerischer, Havriliak-Negami, fractal, and PWC models are presented in Figures S1, S2, S3, S4, and S5, respectively. For all models, the first derivative of the real part of the GP-regressed impedance was close to the derivative of the real part of the exact impedance (see the Bode plots in Figure 2, S1, S2, S3, S4, and S5). As for the first derivative of the real part of the experimental impedance, it had significant variations. The same observations could be made for the first derivative of the imaginary part of the GP-regressed and experimental impedances. We drew the boxplots of OMSE_∂ for the 1×ZARC and separated 2×ZARC models for $\sigma_n^{\text{exp}} = 0.1, 0.2,\ldots, 1.0\ \Omega$ (Figure S8), and for each model and $\sigma_n^{\text{exp}} = 0.1, 0.3, 0.5,$ and $1.0\ \Omega$ (Figure S9). For all models, we observed similar values of OMSE_∂, which increased monotonically with σ_n^{exp}, consistent with intuition. Additionally, we compared OMSE_∂ to the $\text{OMSE}_{\partial,\text{exp}}$ defined in (12b). Based on (7) and (12b), $\text{OMSE}_{\partial,\text{exp}}$ only depends on σ_n^{exp}, like OMSE_{exp}, see Figures S10 and S11. We also found that $\text{OMSE}_\partial \ll \text{OMSE}_{\partial,\text{exp}}$ in every case studied.

3.1.2   Investigation of other Noise Models

In the synthetic experiments considered in Section 3.1.1, the random errors, $\boldsymbol{\varepsilon}$, defined in (7), were drawn such as an independent and identically distributed normal random process with zero mean. In other words, the errors were chosen to be independent of the frequency [29,30,35,41]. As such independence may not necessarily hold experimentally [56], we also considered an experimental noise model with [57,58]

$$\sigma_n^{\text{exp}}(f) \propto |Z_{\text{exact}}(f)| \tag{24}$$

so that the experimental noise was drawn from

$$\boldsymbol{\varepsilon} \sim \mathcal{N}\left(0, \begin{pmatrix} \left(\sigma_n^{\text{exp}}(f_1)\right)^2 & 0 & \cdots & & 0 \\ 0 & \left(\sigma_n^{\text{exp}}(f_2)\right)^2 & & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & & \cdots & 0 & \left(\sigma_n^{\text{exp}}(f_M)\right)^2 \end{pmatrix} \right) \tag{25}$$

with $(f_1, \ldots, f_M)$ being the frequencies probed and all the off-diagonal values are equal to 0. It is worth stressing that a coefficient of proportionality of 0.1 leads to $\frac{1}{M}\sum_{k=1}^{M}\sigma_n^{\text{exp}}(f_k) = 2.47 \ \Omega$ for the 1×ZARC model and log-spaced frequency range from 10^{-2} to 10^6 Hz with 10 ppd. We must note that this averaged value is larger than the noise level of $1.0 \ \Omega$ used in Section 3.1.1.

We then followed the same GP procedure used in Section 3.1.1, and, in particular, we did not modify the matrix $\mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega})$, see (6c). Panels (a) and (b) of Figure S10 show the Bode plots of the GP-regressed and experimental impedances for the 1×ZARC with the noise model (24). The EIS spectrum was closely regressed (see the credible bands close to the mean in panels (a) and (b) of Figure S12. GPs were also able to filter and smooth the exact impedance corrupted by this type of noise. This can be understood graphically (Figure S12), and with the boxplots of OMSE, OMSE_{exp}, OMSE_∂, and $\text{OMSE}_{\partial,\text{exp}}$ (Figure S13) where $\text{OMSE} \ll \text{OMSE}_{\text{exp}}$ and $\text{OMSE}_\partial \ll \text{OMSE}_{\partial,\text{exp}}$.

We also considered an error model such that the real and imaginary parts of the exact impedance were each corrupted differently and independently, see (S10) in Section S1.4 [57,58]. As for the noise model (24), we followed the same procedure as in Section 3.1.1 without modifying the matrix $\mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega})$ in (6c). We obtained the Bode plots of the GP-regressed and experimental impedances, and the Bode plots of the first derivatives of the exact, GP-regressed, and experimental impedances

(Figure S14). The boxplots of OMSE, $\mathrm{OMSE_{exp}}$, $\mathrm{OMSE_{\partial}}$, and $\mathrm{OMSE_{\partial,exp}}$ are shown in Figure S13, and we found $\mathrm{OMSE} \ll \mathrm{OMSE_{exp}}$ and $\mathrm{OMSE_{\partial}} \ll \mathrm{OMSE_{\partial,exp}}$ like for the noise model (24).

3.2 Polarisation Resistance

3.2.1 Validation with Stochastic Experiments

Using the artificial EIS spectra from Section 3.1, we derived μ_{R_p} and σ_{R_p} given in (17a) and (17b). Panels (b), (d), and (f) of Figure 5 show the boxplots of R_p as functions of σ_n^{exp} in the range from $\sigma_n^{\mathrm{exp}} = 0.1, 0.2,\ldots, 1.0\ \Omega$ for the 1×ZARC, separated 2×ZARC, and distant 2×ZARC models, respectively. The predicted median matched well with the exact value of the R_p for the 1×ZARC ($R_\mathrm{p} = 50\ \Omega$), separated 2×ZARC ($R_\mathrm{p} = 100\ \Omega$), and distant 2×ZARC ($R_\mathrm{p} = 70\ \Omega$) models, see the red dashed lines in panels (b), (d), and (f) of Figure 5. As expected, the standard deviations increased with σ_n^{exp}.
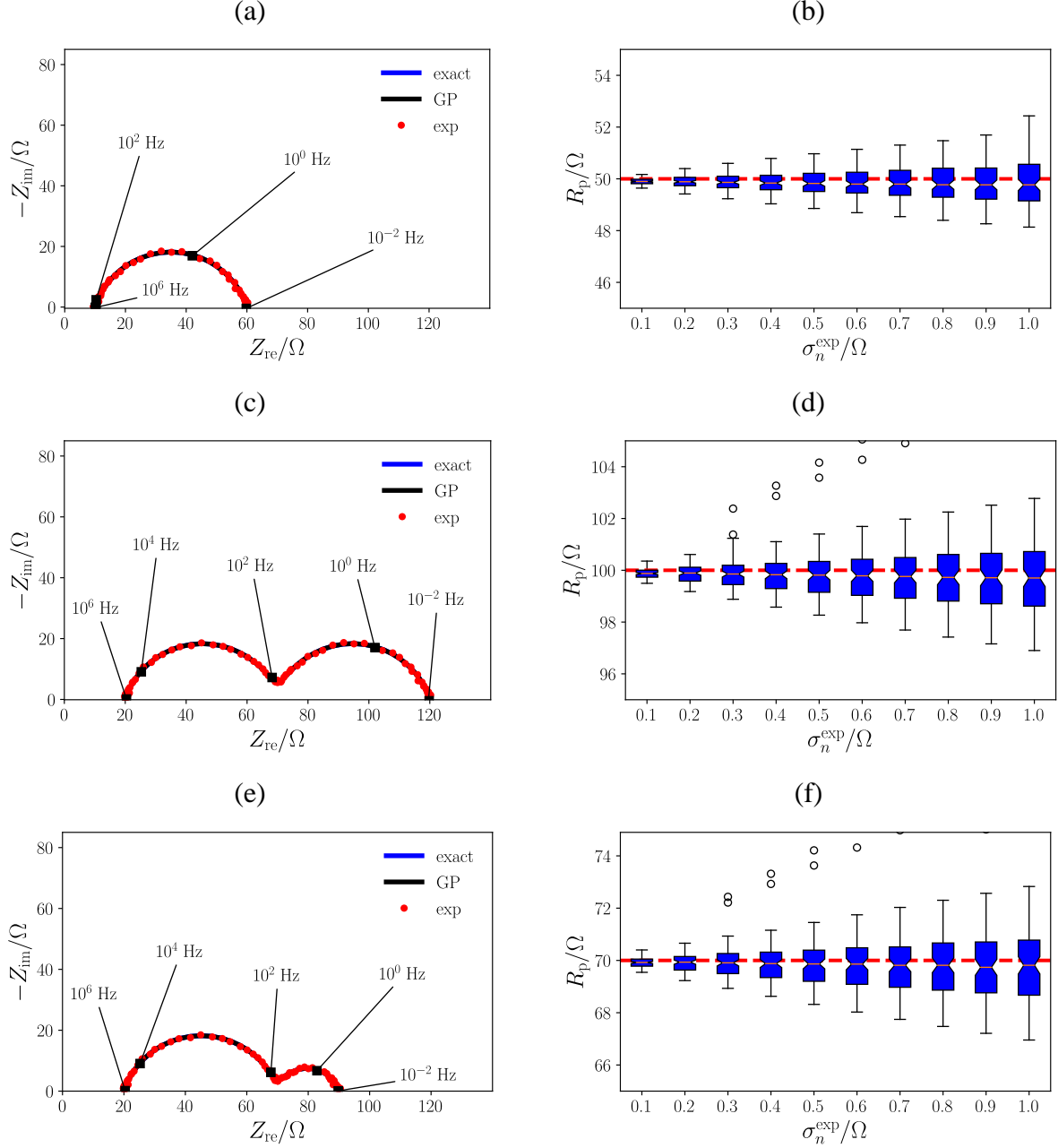
Figure 5. For $\sigma_n^{\text{exp}} = 0.3\ \Omega$, Nyquist plot of a (a) 1×ZARC, (c) separated 2×ZARC, and (e) distant 2×ZARC models; corresponding boxplots of R_{p} as functions of σ_n^{exp} (c), (d), and (f).

3.2.2 Influence of the Cut-off Frequency

It often occurs that the measured EIS data is truncated at low frequencies due to time constraints, or at high frequencies due to experimental setup issues. In such cases, a direct evaluation of the R_{p}

as the diameter of the Nyquist plot may prove either inaccurate or impossible. For these reasons, it is crucial to find out how the set of probed frequencies influences the R_p prediction. To that end, we considered three truncated ranges with 10 ppd, *i.e.*, from 10^{-1} to 10^6 Hz (truncation at low frequencies, 71 frequencies probed), 10^{-2} to 10^2 Hz (truncation at intermediate frequencies, 41 frequencies probed), and 10^{-2} to 10^3 Hz (truncation at high frequencies, 51 frequencies probed). Then, given the 1×ZARC model and $\sigma_n^\mathrm{exp} = 1.0\ \Omega$, we generated 100 synthetic experiments for each frequency range. Subsequently, we applied the same GP methodology as in Section 3.2.1 to estimate μ_{R_p} and σ_{R_p} Figure 6 displays the boxplots of the R_p for each frequency range. The median matched approximately the exact value of the R_p (50 Ω) represented by the red dashed line. Additionally, all three boxes remained small and approximately the same size as the box corresponding to the complete frequency range (rightmost box in Figure 6).
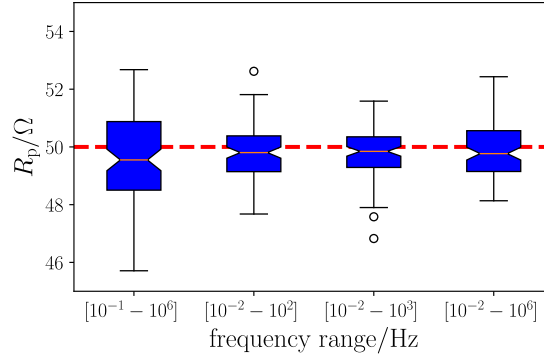
Figure 6. Boxplots of R_p for log-spaced frequencies in the ranges from 10^{-1} to 10^6 Hz (truncation at low frequencies), 10^{-2} to 10^2 Hz (truncation at intermediate frequencies), 10^{-2} to 10^3 Hz (truncation at high frequencies), and 10^{-2} to 10^6 Hz (complete frequency range) with 10 ppd.

3.2.3 Validation with Real Experiments

To complete the previous analysis based on synthetic data, we studied real EIS spectra from one battery, one supercapacitor, and one solar cell.

21

3.2.3.1.1 Rechargeable Lithium-metal Battery

We first computed the R_p of a rechargeable Li-metal battery with a LiFePO$_4$ cathode and a composite polymer electrolyte [59]. The potentiometric EIS data was obtained in the frequency range from 1 Hz to 0.9 MHz using a voltage amplitude of 10 mV. In addition to the GP analysis illustrated in Section 3.1, we regressed the EIS spectrum with an ECM consisting of three ZARCs whose parameters values are given in Table S3. Figure 7 (a) shows the Nyquist plot of the experimental, ECM, and GP-regressed impedances. Only the frequency range from 34 Hz to 58 kHz was used for GP regression and prediction. We found that GPs closely regressed the experimental EIS spectrum (Figure 7 (a) and S15). Moreover, we obtained $R_p = 1.53 \text{ k}\Omega \pm 43.7 \times 10^{-3}$ kΩ, which is consistent with the Nyquist plot and the ECM value of 1.48 kΩ (Figure 7 (a)).
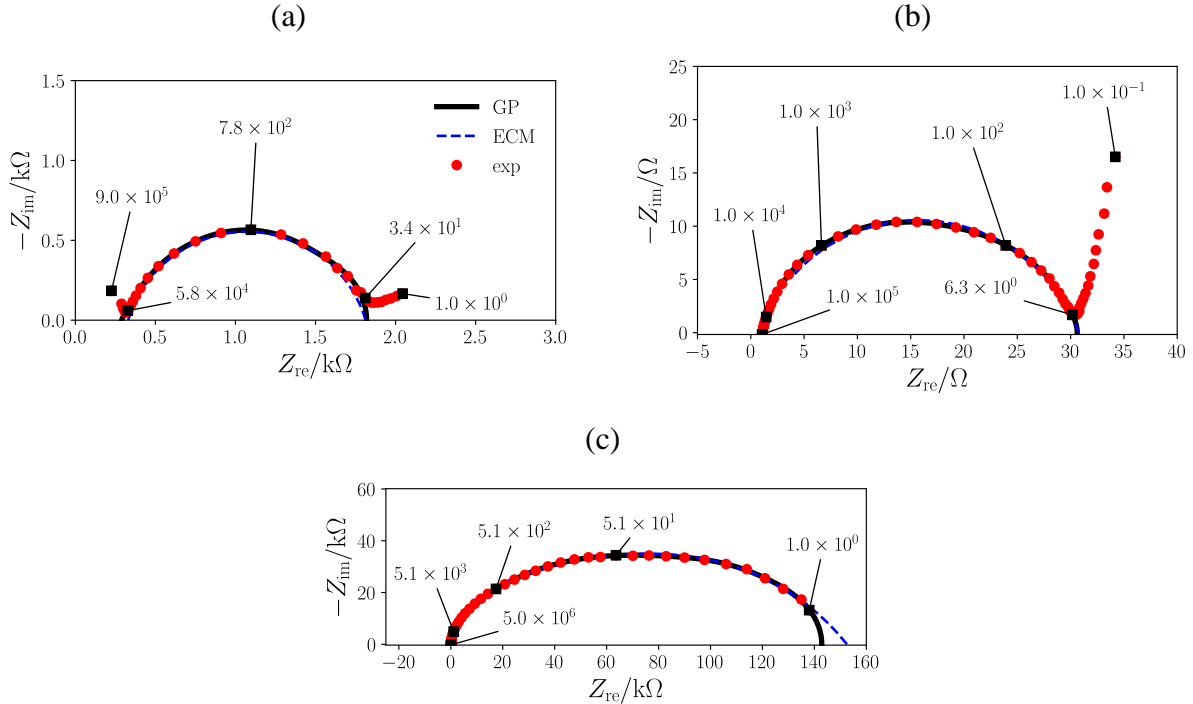
Figure 7. Nyquist plot of the experimental, ECM-fitted, and GP-regressed impedances for (a) a Li-metal battery, (b) a double-layer supercapacitor, and (c) a perovskite solar cell.

3.2.3.2 Electric Double-Layer Supercapacitor

Next, we analyzed the EIS spectrum collected from an electric double-layer supercapacitor with porous carbone electrodes and a tetraethylammonium-in-acetonitrile electrolyte [60]. The EIS spectrum was acquired at frequencies in the 0.1 Hz to 0.1 MHz range with a 5 mV voltage amplitude. Consistent with the electrochemical nature of this system, we observed diffusional features at low frequencies, see Figure 7 (b). We therefore censored frequencies below 6.3 Hz. The parameters values of the 2×ZARC ECM are presented in Table S3. The experimental EIS spectrum was regressed closely by GPs (Figure 7 (b) and S16). Lastly, we computed $R_p = 3.00 \times 10^1 \ \Omega \pm 3.01 \times 10^{-1} \ \Omega$, which is consistent with Figure 7 (b) and the ECM value of $2.998 \times 10^1 \ \Omega$.

3.2.3.3 Perovskite Solar Cell

We also studied EIS data measured from a perovskite solar cell. The cell was made of fluorine-doped tin oxide, the electron-transport layer of SnO_2, the transport-hole layer of Spiro-OMeTAD, and the absorber of $(FAPbI_3)_{0.85}(MAPbBr_3)_{0.15}$ [23]. The EIS spectrum was measured from 1 Hz to 5 MHz in a 1 atm atmosphere with a 100 mV voltage amplitude. The ECM was fitted against a model developed in [61], and the parameter values are given in Table S4. The impedance of this solar cell was regressed over the whole frequency range and was predicted below 1 Hz (Figure 7 (c)). We found that GPs regressed precisely the experimental impedance (Figure 7 (c) and S17). Additionally, we derived $R_p = 1.43 \times 10^2 \ k\Omega \pm 2.27 \times 10^{-1} \ k\Omega$, which is consistent with the Nyquist plot and the ECM value of $1.52 \times 10^2 \ k\Omega$ (Figure 7 (c)).

3.3 Active Learning for EIS

In this section, we apply the acquisition method described in Section 2.2.2. We first must note that the implementation of the ALC was more time consuming compared to that of the ALM as the

integral in (21) needs to be computed. The first, second, and third rows of Figure 8 show σ^2_{avg}, σ^2_{max}, and OMSE as functions of the iteration number, respectively. In that figure, the results shown correspond to the use of the A-criterion for ALM and ALC. Corresponding plots for the D- and E-criteria are shown in Figures S18 and S19. We observed that most curves were discontinuous after 30 iterations, which was a consequence of the hyperparameter value update (see Section 2.2.2). Furthermore, the values of σ^2_{avg}, σ^2_{max}, and OMSE were in most cases lower with the ALC than with the ALM, independent of the criterion used, implying that, consistent with the theory, the ALC achieved a higher accuracy [45]. Additionally, σ^2_{avg} and σ^2_{max} decreased monotonically with the iteration number, eventually reaching a plateau. We also found that all values of σ^2_{avg} were lower than $\sigma^2_{avg,5ppd}$ (panels (a) and (b) of Figure 8, S18 and S19), and in some cases lower than $\sigma^2_{avg,10ppd}$. In other words, this acquisition method can identify frequencies characterized by a better average GP regression than that obtained using a longer 10 ppd experiment. The same conclusion applies for σ^2_{max} (panels (c) and (d) of Figure 8, S18 and S19) and OMSE, whose values remained, in most cases, smaller than $OMSE_{5ppd}$ (panels (e) and (f) of Figure 8, S18 and S19). Interestingly, we also noticed that some values of OMSE were smaller than $OMSE_{10ppd}$ (Figure 8 (f) and Figure S19 (e) and (f)).

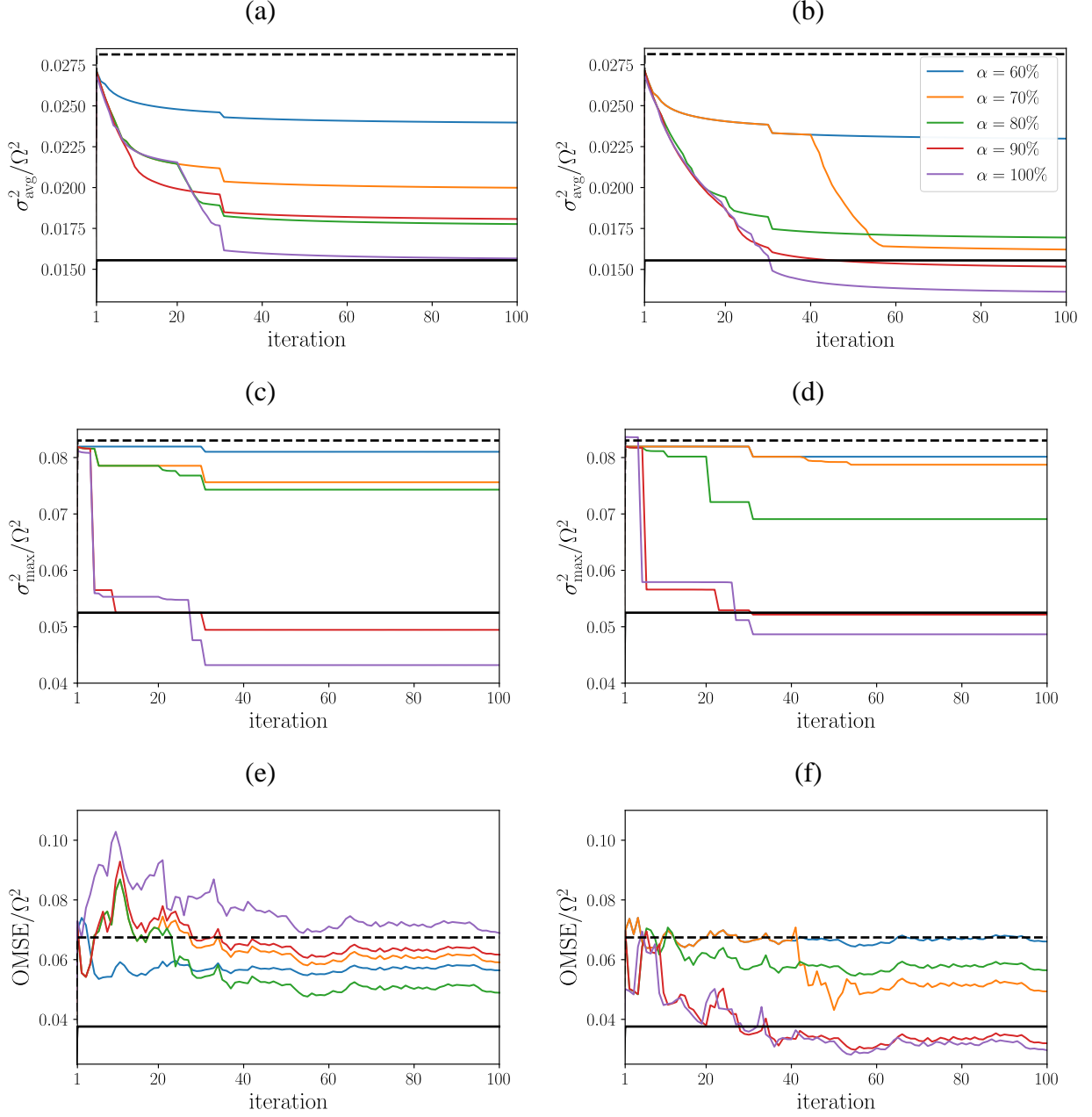Figure 8. (a) σ^2_{avg}, (c) σ^2_{max}, and (e) OMSE as functions of the iteration number for the A-criterion ALM; corresponding values for the A-criterion ALC are shown in panels (b), (d), and (f). The dashed and solid lines on each plot refer to the 5 and 10 ppd cases, respectively.

4 Conclusions

In the field of EIS, GPs have been previously used to recover the impedance from the distribution of relaxation times and to analyze the quality of the EIS spectrum through the Hilbert or Kramers-Kronig transforms. This article proposes three new applications of GPs to augment EIS experiments. We used GPs to regress, filter, and smooth noisy EIS spectra, and to predict EIS spectra and polarization resistances. Synthetic and real EIS data were used to validate the approach. Moreover, we developed a new active learning acquisition framework based on GPs to enable quick and accurate EIS measurements. We envision that this work will open new research and application opportunities for the probabilistic analysis and acquisition of EIS data.

Code Availability

Relevant code is available at https://github.com/ciuccislab/GP-HT.

Author Credit Statement

Baptiste Py: Extension, Development, and Validation of software, Methodology, Theory, Experimental Analysis, Writing – Original Draft, Writing –, Review & Editing. Adeleke Maradesa: Investigation, Writing – Review & Editing. Francesco Ciucci: Software Creation, Conceptualization, Methodology, Theory development, Investigation, Resources, Review & Editing, Funding acquisition, Project administration, Supervision.

Declaration of Competing Interest

None.

Acknowledgments

List of Symbols

$\boldsymbol{\varepsilon}$	Vector of random experimental errors
$\boldsymbol{\mu}$	Mean vector
$\boldsymbol{\Sigma}$	Covariance matrix
$\boldsymbol{\omega}_{5\mathrm{ppd}}$	Vector of log-spaced angular frequencies in the range from $2\pi.10^{-2}$ to $2\pi.10^{6}$ rad/s with 5 ppd
$\mu_{R_{\mathrm{p}}}$	Predicted mean of the polarization resistance
σ_{avg}	GP regression average error
σ_{max}	GP regression maximum error
σ_{BL}	Parameter of the band-limited kernels
σ_{L}	Standard deviation of the inductance
σ_{n}	Standard deviation of the random error
σ_{R}	Standard deviation of the resistance

σ_{SB}	Parameter of the stationary-based kernels
$\sigma_{R_{\text{p}}}$	Predicted standard deviation of the polarization resistance
σ_n^{exp}	Standard deviation of the experimental noise
τ_{min}	Lower band of the BL kernel
τ_{max}	Upper band of the BL kernel
ω	Angular frequency

$\boldsymbol{I}$	Identity matrix
$\mathbf{K}$	Covariance matrix
$\boldsymbol{Z}, \boldsymbol{Z}_*$	Vectors of impedances
f	Frequency
k	Kernel
l_{SB}	Parameter of the stationary-based kernels
M	Number of frequencies
R_{p}	Polarization resistance
Z	Impedance

List of Abbreviations

ALC	Active learning Cohn
ALM	Active learning MacKay
BL	Band-limited
ECM	Equivalent circuit model
EIS	Electrochemical impedance spectroscopy
GP	Gaussian process
OMSE	Out-of-sample mean square error
$OMSE_{\partial}$	Out-of-sample mean square derivative error
PPD	Points per decade
PWC	Piecewise constant
SB	Stationary-based

References

[1]   V. Vivier, M.E. Orazem, Impedance analysis of electrochemical systems, Chem. Rev. 122 (2022) 11131–11168.

[2]   P. Iurilli, C. Brivio, V. Wood, On the use of electrochemical impedance spectroscopy to characterize and model the aging phenomena of lithium-ion batteries: a critical review, J. Power Sources. 505 (2021) 229860.

[3]   Y. Zhang, Y. Katayama, R. Tatara, L. Giordano, Y. Yu, D. Fraggedakis, J.G. Sun, F. Maglia, R. Jung, M.Z. Bazant, Y. Shao-Horn, Revealing electrolyte oxidation *via* carbonate

dehydrogenation on Ni-based oxides in Li-ion batteries by *in situ* Fourier transform infrared spectroscopy, Energy Environ. Sci. 13 (2020) 183–199.

[4]   E. Quattrocchi, T.H. Wan, A. Curcio, S. Pepe, M.B. Effat, F. Ciucci, A general model for the impedance of batteries and supercapacitors: The non-linear distribution of diffusion times, Electrochim. Acta. 324 (2019) 134853.

[5]   P. Navalpotro, M. Anderson, R. Marcilla, J. Palma, Insights into the energy storage mechanism of hybrid supercapacitors with redox electrolytes by electrochemical impedance spectroscopy, Electrochim. Acta. 263 (2018) 110–117.

[6]   P. Kurzweil, M. Shamonin, State-of-charge monitoring by impedance spectroscopy during long-term self-discharge of supercapacitors and lithium-ion batteries, Batteries. 4 (2018) 35.

[7]   A. Sacco, Electrochemical impedance spectroscopy: Fundamentals and application in dye-sensitized solar cells, Renewable and Sustainable Energy Reviews. 79 (2017) 814–829.

[8]   A.R.C. Bredar, A.L. Chown, A.R. Burton, B.H. Farnum, Electrochemical impedance spectroscopy of metal oxide electrodes for energy applications, ACS Appl. Energy Mater. 3 (2020) 66–98.

[9]   A. Omar, M.S. Ali, N. Abd Rahim, Electron transport properties analysis of titanium dioxide dye-sensitized solar cells (TiO2-DSSCs) based natural dyes using electrochemical impedance spectroscopy concept: A review, Solar Energy. 207 (2020) 1088–1121.

[10]   S. Wang, J. Zhang, O. Gharbi, V. Vivier, M. Gao, M. E. Orazem, Electrochemical impedance spectroscopy, Nat. Rev. Methods Primers 1. (2021) 1-21.

[11]   H.S. Magar, R.Y.A. Hassan, A. Mulchandani, Electrochemical impedance spectroscopy (EIS): Principles, construction, and biosensing applications, Sensors. 21 (2021) 6578.

[12]   E. Kozhevnikov, X. Hou, S. Qiao, Y. Zhao, C. Li, W. Tian, Electrical impedance spectroscopy – a potential method for the study and monitoring of a bone critical-size defect healing process treated with bone tissue engineering and regenerative medicine approaches, J. Mater. Chem. B. 4 (2016) 2757–2767.

[13]   K. Krukiewicz, Electrochemical impedance spectroscopy as a versatile tool for the characterization of neural tissue: A mini review, Electrochem. Commun. 116 (2020) 106742.

[14]   M. Grossi, B. Riccò, Electrical impedance spectroscopy (EIS) for biological analysis and food characterization: a review, J. Sens. Sens. Syst. 6 (2017) 303–325.

[15]   A. ter Heijne, D. Liu, M. Sulonen, T. Sleutels, F. Fabregat-Santiago, Quantification of bio-anode capacitance in bioelectrochemical systems using electrochemical impedance spectroscopy, J. Power Sources. 400 (2018) 533–538.

[16]   P. Vadhva, J. Hu, M.J. Johnson, R. Stocker, M. Braglia, D.J.L. Brett, A.J.E. Rettie, Electrochemical impedance spectroscopy for all-solid-state batteries: Theory, methods and future outlook, Chem. Electrochem. 8 (2021) 1930–1947.

[17]   I. Pivac, B. Šimić, F. Barbir, Experimental diagnostics and modeling of inductive phenomena at low frequencies in impedance spectra of proton exchange membrane fuel cells, J. Power Sources. 365 (2017) 240–248.

[18]   T. Osaka, D. Mukoyama, H. Nara, Review—Development of diagnostic process for commercially available batteries, especially lithium ion battery, by electrochemical impedance spectroscopy, J. Electrochem. Soc. 162 (2015) A2529–A2537.

[19]   Q. Meyer, Y. Zeng, C. Zhao, Electrochemical impedance spectroscopy of catalyst and carbon degradations in proton exchange membrane fuel cells, J. Power Sources. 437 (2019) 226922.

[20]   W. Zhang, J. Ma, P. Wang, Z. Wang, F. Shi, H. Liu, Investigations on the interfacial capacitance and the diffusion boundary layer thickness of ion exchange membrane using electrochemical impedance spectroscopy, J. Membrane Sci. 502 (2016) 37–47.

[21]   T. Pajkossy, R. Jurczakowski, Electrochemical impedance spectroscopy in interfacial studies, Curr. Opin. Electrochem. 1 (2017) 53–58.

[22]   D.A. Harrington, P. van den Driessche, Mechanism and equivalent circuits in electrochemical impedance spectroscopy, Electrochim. Acta. 56 (2011) 8005–8013.

[23]   E. Quattrocchi, T.H. Wan, A. Belotti, D. Kim, S. Pepe, S.V. Kalinin, M. Ahmadi, F. Ciucci, The deep-DRT: A deep neural network approach to deconvolve the distribution of relaxation times from multidimensional electrochemical impedance spectroscopy data, Electrochim. Acta. 392 (2021) 139010.

[24]   S. Dierickx, A. Weber, E. Ivers-Tiffée, How the distribution of relaxation times enhances complex equivalent circuit models for fuel cells, Electrochim. Acta. 355 (2020) 136764.

[25] J. Huang, M. Papac, R. O'Hayre, Towards robust autonomous impedance spectroscopy analysis: A calibrated hierarchical Bayesian approach for electrochemical impedance spectroscopy (EIS) inversion, Electrochim. Acta. 367 (2021) 137493.

[26] F. Ciucci, Modeling electrochemical impedance spectroscopy, Curr. Opin. Electrochem. 13 (2019) 132–139.

[27] Q. Meyer, I. Pivac, F. Barbir, C. Zhao, Detection of oxygen starvation during carbon corrosion in proton exchange membrane fuel cells using low-frequency electrochemical impedance spectroscopy, J. Power Sources. 470 (2020) 228285.

[28] A. Lasia, Electrochemical impedance spectroscopy and its applications, Modern Aspects Electrochem. (2002) 143-248.

[29] C.E. Rasmussen, C.K.I. Williams, Gaussian processes for machine learning, MIT Press, Cambridge, Mass, 2006.

[30] A. Maradesa, B. Py, E. Quattrocchi, F. Ciucci, The probabilistic deconvolution of the distribution of relaxation times with finite Gaussian processes, Electrochim. Acta. 413 (2022) 140119.

[31] S. Särkkä, Bayesian filtering and smoothing. No. 3., Cambridge University Press (2013).

[32] E. Schulz, M. Speekenbrink, A. Krause, A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions, J. Mathematical Psychology. 85 (2018) 1–16.

[33] D.A. Cohn, Z. Ghahramani, M.I. Jordan, Active learning with statistical models, J. Artif. Intell. Research 4 (1996) 129-145.

[34]   D.J.C. MacKay, Information-based objective functions for active data selection, Neural Computation. 4 (1992) 590–604.

[35]   J. Liu, F. Ciucci, The Gaussian process distribution of relaxation times: A machine learning tool for the analysis and prediction of electrochemical impedance spectroscopy data, Electrochim. Acta. 331 (2020) 135316.

[36]   Y. Lu, C.-Z. Zhao, J.-Q. Huang, Q. Zhang, The timescale identification decoupling complicated kinetic processes in lithium batteries, Joule. 6 (2022) 1172–1198.

[37]   E. IVERS-TIFFÉE, A. Weber, Evaluation of electrochemical impedance spectra by the distribution of relaxation times, J. Ceram. Soc. Japan. 125 (2017) 193–201.

[38]   B.A. Boukamp, A. Rolle, Analysis and application of distribution of relaxation times in solid state ionics, Solid State Ionics. 302 (2017) 12–18.

[39]   M. Heinzmann, A. Weber, E. Ivers-Tiffée, Advanced impedance study of polymer electrolyte membrane single cells by means of distribution of relaxation times, J. Power Sources. 402 (2018) 24–33.

[40]   J. Liu, T.H. Wan, F. Ciucci, A Bayesian view on the Hilbert transform and the Kramers-Kronig transform of electrochemical impedance data: Probabilistic estimates and quality scores, Electrochim. Acta. 357 (2020) 136864.

[41]   F. Ciucci, The Gaussian process Hilbert transform (GP-HT): Testing the consistency of electrochemical impedance spectroscopy data, J. Electrochem. Soc. 167 (2020) 126503.

[42]   F. W. King, Hilbert Transforms, Cambridge University Press, Cambridge, 2009.

[43]   B.A. Boukamp, Distribution (function) of relaxation times, successor to complex nonlinear least squares analysis of electrochemical impedance spectroscopy?, J. Phys. Energy. 2 (2020) 042001.

[44]   D. Malevich, E. Halliop, B.A. Peppley, J.G. Pharoah, K. Karan, Investigation of charge-transfer and mass-transport esistances in PEMFCs with microporous layer using electrochemical impedance spectroscopy, J. Electrochem. Soc. 156 (2009) B216.

[45]   R.B. Gramacy, Surrogates: Gaussian process modeling, design, and optimization for the applied sciences, Chapman and Hall/CRC (2020).

[46]   D. A. Cohn, Neural network exploration using optimal experiment design, Advances Neural Info. Process. Systems. 6 (1993).

[47]   F. Ciucci, T. Carraro, W.C. Chueh, W. Lai, Reducing error and measurement time in impedance spectroscopy using model based optimal experimental design, Electrochim. Acta. 56 (2011) 5416–5434.

[48]   S. Effendy, J. Song, M.Z. Bazant, Analysis, design, and generalization of electrochemical impedance spectroscopy (EIS) inversion algorithms, J. Electrochem. Soc. 167 (2020) 106508.

[49]   T.H. Wan, M. Saccoccio, C. Chen, F. Ciucci, Influence of the discretization methods on the distribution of relaxation times deconvolution: Implementing radial basis functions with DRTtools, Electrochim. Acta. 184 (2015) 483–499.

[50]   M.B. Effat, F. Ciucci, Bayesian and hierarchical Bayesian based regularization for deconvolving the distribution of relaxation times from electrochemical impedance spectroscopy data, Electrochim. Acta. 247 (2017) 1117–1129.

[51]   Higham, Nicholas J., Computing a nearest symmetric positive semidefinite matrix, Linear Algebra and Its Applications. (1988).

[52]   T. Holsclaw, B. Sansó, H.K.H. Lee, K. Heitmann, S. Habib, D. Higdon, U. Alam, Gaussian process modeling of derivative curves, Technometrics. 55 (2013) 57–67.

[53]   D. Klotz, Negative capacitance or inductive loop? – A general assessment of a common low frequency impedance feature, Electrochem. Commun. 98 (2019) 58–62.

[54]   M. Saccoccio, T.H. Wan, C. Chen, F. Ciucci, Optimal regularization in distribution of relaxation times applied to electrochemical impedance spectroscopy: Ridge and Lasso regression methods - A theoretical and experimental study, Electrochim. Acta. 147 (2014) 470–482.

[55]   F. Ciucci, C. Chen, Analysis of electrochemical impedance spectroscopy data using the distribution of relaxation times: A Bayesian and hierarchical Bayesian approach, Electrochim. Acta. 167 (2015) 439–454.

[56]   D.M. Morgan, S.G. Weber, Noise and signal-to-noise ratio in electrochemical detectors, Anal. Chem. 56 (1984) 2560–2567.

[57]   M.E. Orazem, P. Agarwal, C. Deslouis, B. Tribollet, Application of measurement models to electrohydrodynamic impedance spectroscopy, J. Electrochem. Soc. 143 (1996) 948–960.

[58]   P. Agarwal, O.D. Crisalle, M.E. Orazem, L.H. Garcia-Rubio, Application of measurement models to impedance spectroscopy: II . Determination of the stochastic contribution to the error tructure, J. Electrochem. Soc. 142 (1995) 4149–4158.

[59]   Z. Dai, J. Yu, J. Liu, R. Liu, Q. Sun, D. Chen, F. Ciucci, Highly conductive and nonflammable composite polymer electrolytes for rechargeable quasi-solid-state Li-metal batteries, J. Power Sources. 464 (2020) 228182.

[60]   H.D. Yoo, J.H. Jang, J.H. Ryu, Y. Park, S.M. Oh, Impedance analysis of porous carbon electrodes to predict rate capability of electric double-layer capacitors, J. Power Sources. 267 (2014) 411–420.

[61]   D. Kim, E.S. Muckley, N. Creange, T.H. Wan, M.H. Ann, E. Quattrocchi, R.K. Vasudevan, J.H. Kim, F. Ciucci, I.N. Ivanov, S.V. Kalinin, M. Ahmadi, Exploring transport behavior in hybrid perovskites solar cells via machine learning analysis of environmental-dependent impedance spectroscopy, Adv. Sci. 8 (2021) 2002510.

Supplementary Information

Gaussian Processes for the Analysis of Electrochemical Impedance Spectroscopy Data: Prediction, Filtering, and Active Learning

Baptiste Py[a], Adeleke Maradesa[a], Francesco Ciucci[a,b,c,d,*]

[a] Department of Mechanical and Aerospace Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China

[b] Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China

[c] HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, China

[d] HKUST Energy Institute, The Hong Kong University of Science and Technology, Hong Kong SAR, China

[*]Corresponding author: francesco.ciucci@ust.hk

Phone: +852 2358 7187

Fax: +852 2358 1543

1 Theoretical Background

1.1 Gaussian Processes

We follow the notations from the main manuscript.

Assuming that $h(\omega)$ is a GP, $h(\boldsymbol{\omega}) = \big(h(\omega_1), h(\omega_2), \dots, h(\omega_M)\big)^{\mathsf{T}}$ with $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_M)^{\mathsf{T}}$ is a multivariate normal random variable with mean $m(\boldsymbol{\omega}) = \mathbb{E}[h(\boldsymbol{\omega})]$ and covariance matrix $\mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega}) = k(\boldsymbol{\omega}, \boldsymbol{\omega}) = \big(\big(k(\omega_n, \omega_m)\big)\big)$ for $n, m = 1, 2, \dots, M$, *i.e.*, in shorthand notation

$$h(\boldsymbol{\omega}) \sim \mathcal{N}\big(m(\boldsymbol{\omega}), \mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega})\big) \tag{S1}$$

with $\mathbb{E}[\cdot]$ the expectation [1]. Many options exist to formulate the kernel k [1,2].

We now take an experimental dataset $\{(\omega_k, y_k) \mid k = 1, 2, \dots, M\}$ that verifies $m(\boldsymbol{\omega}) = \mathbf{0}$ and $y_k = h(\omega_k) + \varepsilon$ for $k = 1, 2, \dots, M$ with $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$ a white noise of standard deviation σ_n. Given (S1) and that $h(\omega_*)$ and $\boldsymbol{y} = (y_1, y_2, \dots, y_N)^{\mathsf{T}}$ follow a multivariate normal distribution, we can now predict $h(\omega_*)$ at an unknown angular frequency ω_* [1]

$$\begin{pmatrix} h(\omega_*) \\ \boldsymbol{y} \end{pmatrix} \sim \mathcal{N}\left(0, \begin{pmatrix} k(\omega_*, \omega_*) & \mathbf{K}(\omega_*, \boldsymbol{\omega}) \\ \mathbf{K}^{\mathsf{T}}(\omega_*, \boldsymbol{\omega}) & \mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega}) + \sigma_n^2 \boldsymbol{I} \end{pmatrix} \right) \tag{S2}$$

with $\mathbf{K}(\omega_*, \boldsymbol{\omega}) = \big(k(\omega_*, \omega_1), \ k(\omega_*, \omega_2), \dots, \ k(\omega_*, \omega_M)\big)$ and $\boldsymbol{I}$ the $M \times M$ identity matrix. The distribution of $h(\omega_*)$ given the knowledge of the experimental values $\boldsymbol{y}$ is also multivariate normal [1]

$$h(\omega_*) \mid \boldsymbol{y} \sim \mathcal{N}(\mu_*, \sigma_*^2) \tag{S3a}$$

$$\mu_*(\boldsymbol{\omega}) = \mathbf{K}(\omega_*, \boldsymbol{\omega})(\mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega}) + \sigma_n^2 I)^{-1} \boldsymbol{y} \tag{S3b}$$

$$\sigma_*^2(\boldsymbol{\omega}) = k(\omega_*, \omega_*) - \mathbf{K}(\omega_*, \boldsymbol{\omega})(\mathbf{K}(\boldsymbol{\omega}, \boldsymbol{\omega}) + \sigma_n^2 I)^{-1} \mathbf{K}^{\mathsf{T}}(\omega_*, \boldsymbol{\omega}) \tag{S3c}$$

1.2 Kernels

As explained, each kernel is made of a band-limited (BL) and a stationary-based (SB) contribution [3]

$$k_{re,BL}(\omega, \omega') = \frac{\psi(\omega, \omega', \tau_{max}) - \psi(\omega, \omega', \tau_{min})}{\omega^2 - \omega'^2} \sigma_{BL}^2 \tag{S4a}$$

$$k_{re,SB}(\omega, \omega') = 2l_{SB}^2 \sigma_{SB}^2 \left(\frac{1}{2l_{SB}^2 + (\omega - \omega')^2} + \frac{1}{2l_{SB}^2 + (\omega + \omega')^2} \right) \tag{S4b}$$

$$k_{re-im,BL}(\omega, \omega') = -\frac{\omega'}{2(\omega^2 - \omega'^2)} \left(\ln\left(\frac{1 + \omega^2 \tau_{max}^2}{1 + \omega^2 \tau_{min}^2} \right) - \ln\left(\frac{1 + \omega'^2 \tau_{max}^2}{1 + \omega'^2 \tau_{min}^2} \right) \right) \sigma_{BL}^2 \tag{S4c}$$

$$k_{re-im,SB}(\omega, \omega') = -\sqrt{2}l_{SB}\sigma_{SB}^2 \left(\frac{(\omega' - \omega)}{2l_{SB}^2 + (\omega' - \omega)^2} + \frac{(\omega' + \omega)}{2l_{SB}^2 + (\omega' + \omega)^2} \right) \tag{S4d}$$

$$k_{im,BL}(\omega, \omega') = \frac{\zeta(\omega, \omega', \tau_{max}) - \zeta(\omega, \omega', \tau_{min})}{\omega^2 - \omega'^2} \sigma_{BL}^2 \tag{S4e}$$

$$k_{im,SB}(\omega, \omega') = 2l_{SB}^2 \sigma_{SB}^2 \left(\frac{1}{2l_{SB}^2 + (\omega - \omega')^2} - \frac{1}{2l_{SB}^2 + (\omega + \omega')^2} \right) \tag{S4f}$$

with $\psi(\omega, \omega', \tau) = \omega \arctan(\tau\omega) - \omega' \arctan(\tau\omega')$ and $\zeta(\omega, \omega', \tau) = \omega' \arctan(\tau\omega) - \omega \arctan(\tau\omega')$.

1.3 Linear Combination of two Jointly Gaussian Random Variables

We consider two jointly Gaussian distributions X_1 and X_2, and two scalars α_1 and α_2. Let us define the random variable $Y = \alpha_1 X_1 + \alpha_2 X_2$, and the two vectors $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ and $\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$.

Given three scalars u_1, u_2, and v, the characteristic function $\Phi_{\mathbf{X}}$ of $\mathbf{X}$ evaluated at $\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and that Φ_Y of Y at v are

$$\Phi_{\mathbf{X}}(\boldsymbol{u}) = \mathbb{E}\left(e^{i\boldsymbol{u}^T\mathbf{X}}\right) = \mathbb{E}\left(e^{i(u_1 X_1 + u_2 X_2)}\right) \tag{S5a}$$

$$\Phi_Y(v) = \mathbb{E}\big(e^{ivY}\big) = \mathbb{E}\big(e^{iv(\alpha_1 X_1 + \alpha_2 X_2)}\big) \tag{S5b}$$

By identifying (S5a) and (S5b), the following equation holds

$$\Phi_Y(v) = \Phi_X(v\boldsymbol{\alpha}) \tag{S6}$$

Additionally, since X_1 and X_2 are jointly Gaussian distributed, we can write $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$ with a mean $\boldsymbol{\mu}$ and a covariance matrix $\mathbf{C}$. Consequently, the characteristic function Φ_X of $\mathbf{X}$ is [4]

$$\Phi_X(\boldsymbol{u}) = \mathbb{E}\left(\exp\left(i\boldsymbol{u}^T\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{u}^T\mathbf{C}\boldsymbol{u}\right)\right) \tag{S7}$$

(S6) and (S7) lead to

$$\Phi_Y(v) = \mathbb{E}\left(\exp\left(iv\boldsymbol{\alpha}^T\boldsymbol{\mu} - \frac{1}{2}v^2\boldsymbol{\alpha}^T\mathbf{C}\boldsymbol{\alpha}\right)\right) \tag{S8}$$

The mean μ_Y and standard deviation σ_Y of the random variable Y are then derived by identification between (S7) and (S8),

$$\mu_Y = \boldsymbol{\alpha}^T\boldsymbol{\mu} \tag{S9a}$$

$$\sigma_Y = \boldsymbol{\alpha}^T\mathbf{C}\boldsymbol{\alpha} \tag{S9b}$$

1.4 Modulus-dependent Noise Model

Besides the white noise and the noise model described in (24) of the main manuscript, we also studied the following noise model [5,6]

$$\begin{cases} \sigma_n^{\mathrm{exp,re}}(f) \propto \left|Z_{\mathrm{exact,re}}(f)\right| \\ \sigma_n^{\mathrm{exp,im}}(f) \propto \left|Z_{\mathrm{exact,im}}(f)\right| \end{cases} \tag{S10}$$

where $\sigma_n^{\mathrm{exp,re}}$ (respectively $\sigma_n^{\mathrm{exp,im}}$) is the standard deviation of the experimental noise corrupting the real part $Z_{\mathrm{exact,re}}$ (respectively the imaginary part $Z_{\mathrm{exact,im}}$) of the exact impedance. In this case, the real $\boldsymbol{\varepsilon}_{\mathrm{re}}$ and imaginary $\boldsymbol{\varepsilon}_{\mathrm{im}}$ parts of the noise $\boldsymbol{\varepsilon}$ were distributed according to

$$\boldsymbol{\varepsilon}_{\mathrm{re}} \sim \mathcal{N}\left(0, \begin{pmatrix} \left(\sigma_n^{\mathrm{exp,re}}(f_1)\right)^2 & 0 & \cdots & & 0 \\ 0 & \left(\sigma_n^{\mathrm{exp,re}}(f_2)\right)^2 & & & \vdots \\ \vdots & & \ddots & & 0 \\ 0 & & \cdots & 0 & \left(\sigma_n^{\mathrm{exp,re}}(f_M)\right)^2 \end{pmatrix}\right) \qquad \text{(S11a)}$$

$$\boldsymbol{\varepsilon}_{\mathrm{im}} \sim \mathcal{N}\left(0, \begin{pmatrix} \left(\sigma_n^{\mathrm{exp,im}}(f_1)\right)^2 & 0 & \cdots & & 0 \\ 0 & \left(\sigma_n^{\mathrm{exp,im}}(f_2)\right)^2 & & & \vdots \\ \vdots & & \ddots & & 0 \\ 0 & & \cdots & 0 & \left(\sigma_n^{\mathrm{exp,im}}(f_M)\right)^2 \end{pmatrix}\right) \qquad \text{(S11b)}$$

In this case, we found that a coefficient of proportionality of 0.1 leads to an average of $\sigma_n^{\mathrm{exp,re}}$ and $\sigma_n^{\mathrm{exp,im}}$ of 2.39 Ω and 0.42 Ω for the log-spaced frequency range $10^{-2} - 10^6$ Hz with 10 ppd and the 1×ZARC model.

S2. Tables

The expression of each covariance matrix introduced in (6) of the main manuscript is detailed in Table S1.

Table S1 – Expressions of the covariance matrices in the Section 2.1.1.

Covariance matrix	Expression
$\mathbf{k}_{\mathrm{re}}(\omega_*, \boldsymbol{\omega})$	$\left(k_{\mathrm{re}}(\omega_*, \omega_1),\ k_{\mathrm{re}}(\omega_*, \omega_2),\ \ldots,\ k_{\mathrm{re}}(\omega_*, \omega_M)\right)$
$\mathbf{k}_{\mathrm{re-im}}(\omega_*, \boldsymbol{\omega})$	$\left(k_{\mathrm{re-im}}(\omega_*, \omega_1),\ k_{\mathrm{re-im}}(\omega_*, \omega_2),\ \ldots,\ k_{\mathrm{re-im}}(\omega_*, \omega_M)\right)$
$\mathbf{k}_{\mathrm{im-re}}(\omega_*, \boldsymbol{\omega})$	$\left(k_{\mathrm{im-re}}(\omega_*, \omega_1),\ k_{\mathrm{im-re}}(\omega_*, \omega_2),\ \ldots,\ k_{\mathrm{im-re}}(\omega_*, \omega_M)\right)$
$\mathbf{k}_{\mathrm{im}}(\omega_*, \boldsymbol{\omega})$	$\left(k_{\mathrm{im}}(\omega_*, \omega_1),\ k_{\mathrm{im}}(\omega_*, \omega_2),\ \ldots,\ k_{\mathrm{im}}(\omega_*, \omega_M)\right)$
$\mathbf{k}_{\mathrm{re}}(\boldsymbol{\omega}, \boldsymbol{\omega})$	$\left(\left(k_{\mathrm{re}}(\omega_n, \omega_m)\right)\right)$ for $n, m = 1, 2, \ldots, M$
$\mathbf{k}_{\mathrm{re-im}}(\boldsymbol{\omega}, \boldsymbol{\omega})$	$\left(\left(k_{\mathrm{re-im}}(\omega_n, \omega_m)\right)\right)$ for $n, m = 1, 2, \ldots, M$
$\mathbf{k}_{\mathrm{im}}(\boldsymbol{\omega}, \boldsymbol{\omega})$	$\left(\left(k_{\mathrm{im}}(\omega_n, \omega_m)\right)\right)$ for $n, m = 1, 2, \ldots, M$

with $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_M)$.

Table S2 – Numerical values for the parameters of each model used to generate synthetic EIS spectra.

Model	Parameters	Model	Parameters
1×ZARC with inductance	$R_\infty = 10\ \Omega$ $R_{ct} = 50\ \Omega$ $\tau_0 = 0.1\ s$ $\phi = 0.8$ $L_0 = 5 \times 10^{-6}\ H$	Separated 2×ZARC	$R_{\infty,1} = 10\ \Omega$ $R_{ct,1} = 50\ \Omega$ $\tau_1 = 0.1\ s$ $\phi_1 = 0.8$ $R_{\infty,2} = 10\ \Omega$ $R_{ct,2} = 50\ \Omega$ $\tau_2 = 10^{-4}\ s$ $\phi_2 = 0.8$
Distant 2×ZARC	$R_{\infty,1} = 10\ \Omega$ $R_{ct,1} = 20\ \Omega$ $\tau_1 = 0.1\ s$ $\phi_1 = 0.8$ $R_{\infty,2} = 10\ \Omega$ $R_{ct,2} = 50\ \Omega$ $\tau_2 = 10^{-4}\ s$ $\phi_2 = 0.8$	Hook	$R_{\infty,1} = 5\ \Omega$ $R_{ct,1} = 50\ \Omega$ $\tau_1 = 10^{-4}\ s$ $\phi_1 = 0.8$ $R_{\infty,2} = 5\ \Omega$ $R_{ct,2} = -10\ \Omega$ $\tau_2 = 10^{-1}\ s$ $\phi_2 = 0.8$
Gerischer	$R_\infty = 10\ \Omega$ $R_{ct} = 50\ \Omega$ $\tau_0 = 10^{-2}\ s$	Havriliak-Negami	$R_\infty = 10\ \Omega$ $R_{ct} = 50\ \Omega$ $\tau_0 = 10^{-2}\ s$ $\phi = 0.7$ $\psi = 0.8$
PWC	$R_\infty = 10\ \Omega$ $R_{ct} = 50\ \Omega$ $\tau_1 = 10^{-1}\ s$ $\tau_2 = 10^{-4}\ s$	Fractal	$R_\infty = 10\ \Omega$ $R_{ct} = 50\ \Omega$ $\tau_0 = 10^{-2}\ s$ $\phi = 0.6$

Table S3 – ECM parameters obtained by fitting the EIS spectrum of the rechargeable Li-metal battery and the double-layer supercapacitor.

Parameter	Li-metal battery	Double-layer supercapacitor
R_∞	3.96×10^{-60} Ω	1.00×10^0 Ω
$R_{ct,1}$	9.79×10^2 Ω	3.00×10^1 Ω
$R_{ct,2}$	1.48×10^3 Ω	1.00×10^3 Ω
$R_{ct,3}$	3.15×10^2 Ω	-
τ_1	1.03×10^0 s	5.05×10^{-4} s
τ_2	1.91×10^{-4} s	5.00×10^3 s
τ_3	1.05×10^{-7} s	-
ϕ_1	5.00×10^{-1}	7.90×10^{-1}
ϕ_2	8.17×10^{-1}	9.00×10^{-1}
ϕ_3	1.00×10^0	-
L_0	2.51×10^{-26} H	1.61×10^{-6} H

Table S4 – ECM parameters obtained by fitting the EIS spectrum of the perovskite solar cell.

Parameter	Value
R_s	7.11×10^{-5} Ω
R_1	1.07×10^5 Ω
R_2	4.60×10^4 Ω
C_g	1.66×10^{-9} F
C_s	9.65×10^{-12} F
ϕ_g	8.70×10^{-1}
ϕ_s	5.60×10^{-1}
L_0	7.08×10^{-6} H

S3. Figures

(a)

(b)

(c)

(d)
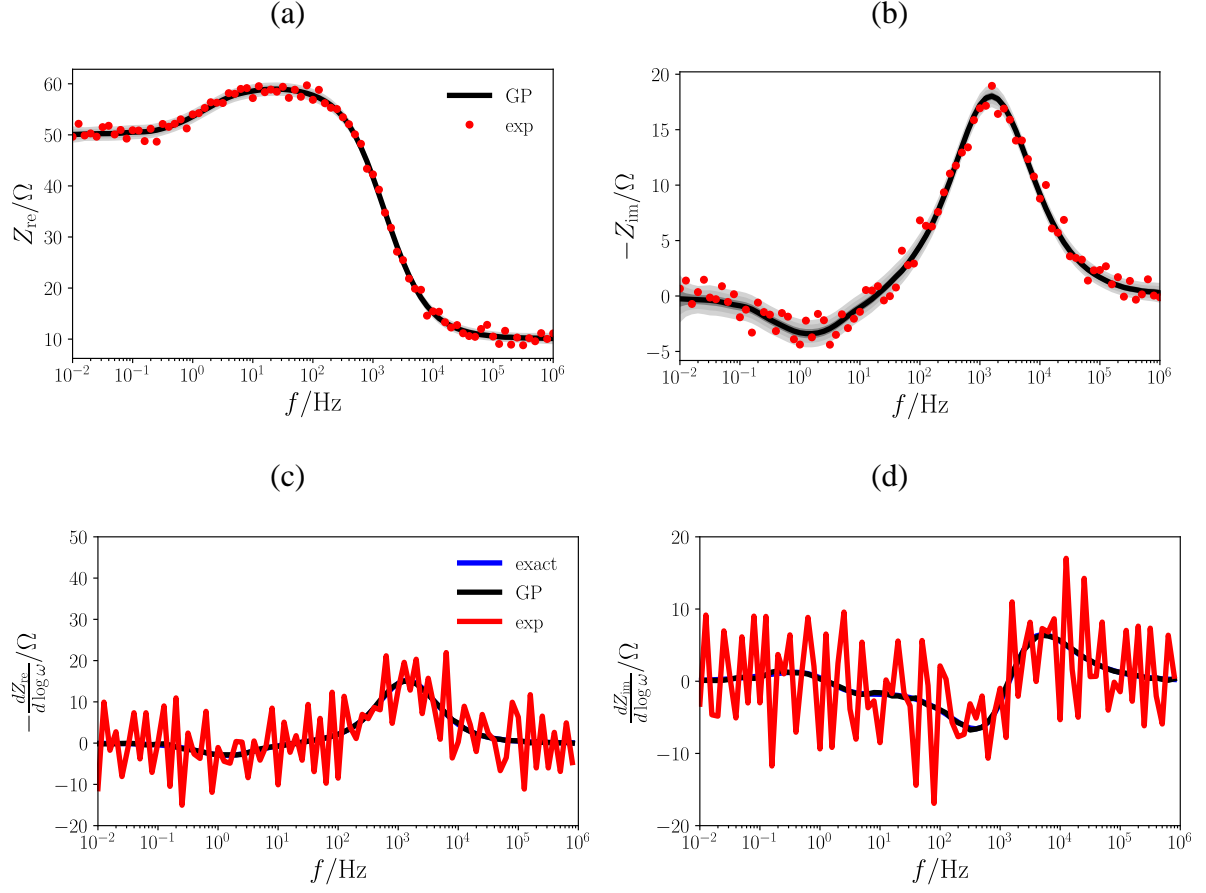
Figure S1. For $\sigma_n^{\mathrm{exp}} = 1.0\ \Omega$ and the hook model, Bode plot of the (a) real and (b) imaginary part of the GP-regressed and experimental impedances; corresponding Bode plots of the first derivative of the exact, GP-regressed, and experimental impedances (c) and (d).
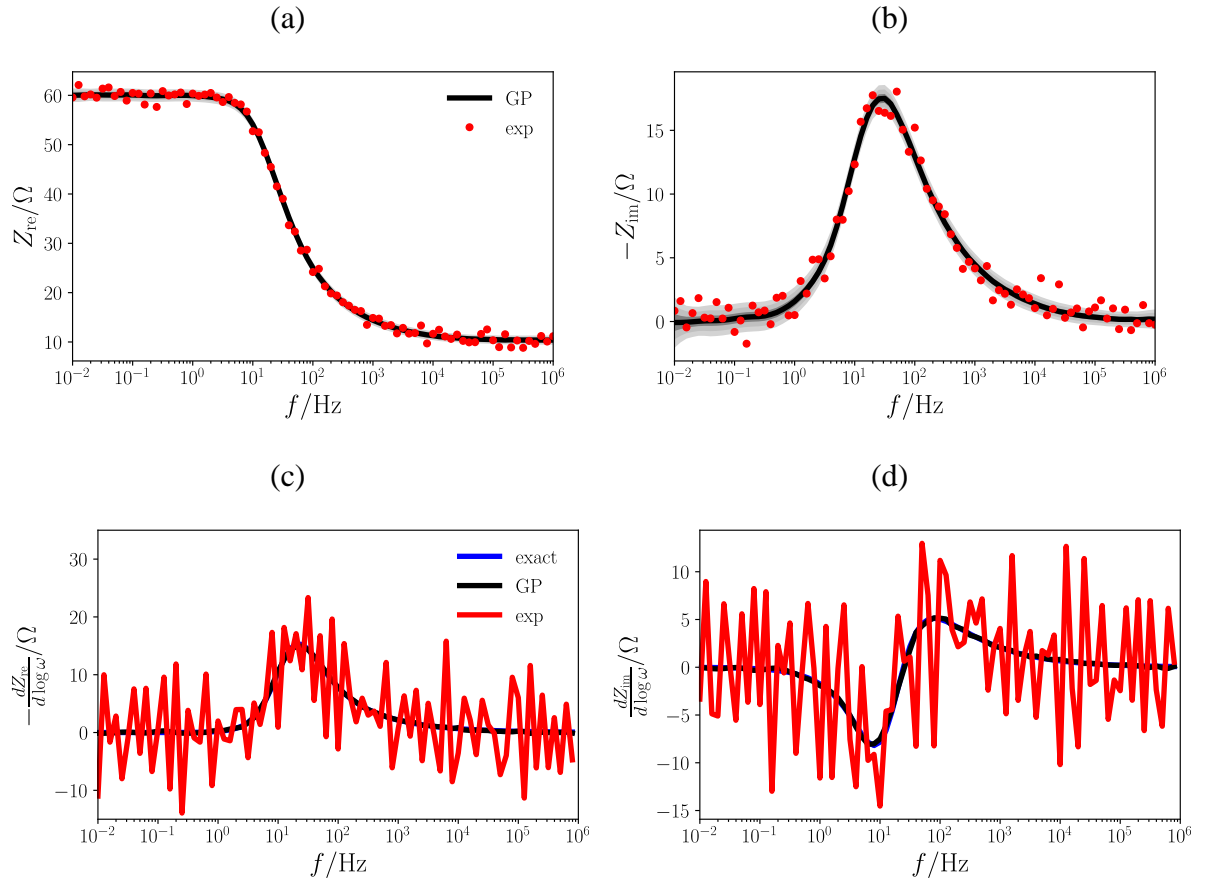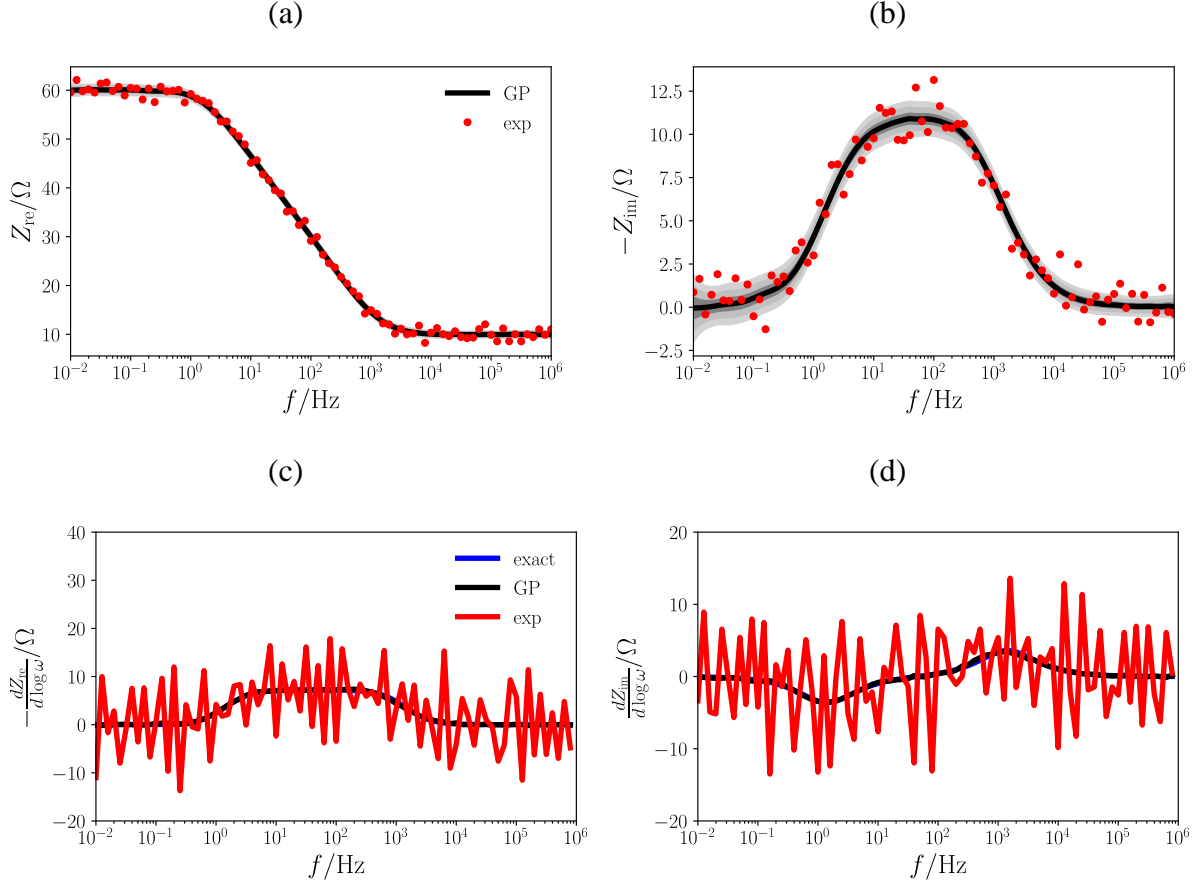
Figure S2. For $\sigma_n^{\text{exp}} = 1.0\,\Omega$ and the Gerischer model, Bode plot of the (a) real and (b) imaginary part of the GP-regressed and experimental impedances; corresponding Bode plots of the first derivative of the exact, GP-regressed, and experimental impedances (c) and (d).

(a) (b)

(c) (d)
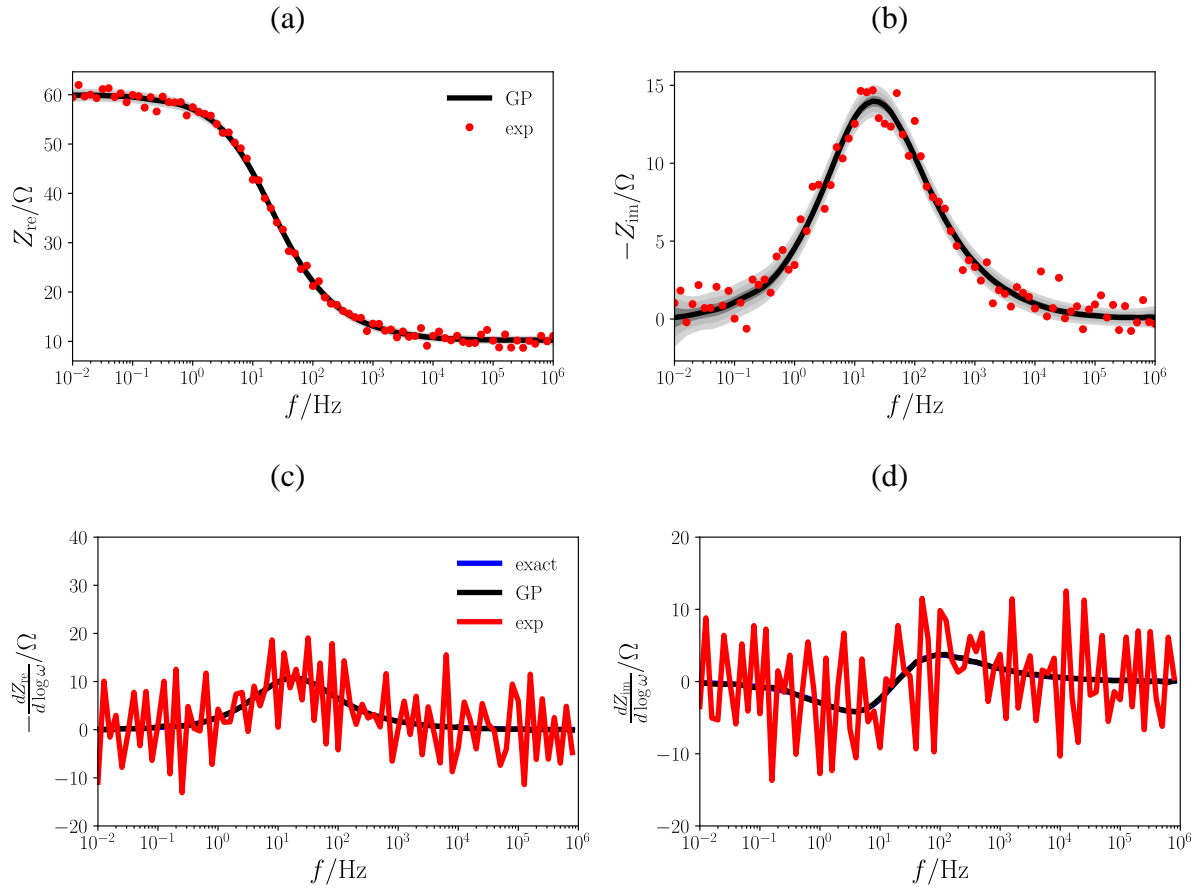
Figure S3. For $\sigma_n^{\mathrm{exp}} = 1.0\ \Omega$ and the Havriliak-Negami model, Bode plot of the (a) real and (b) imaginary part of the GP-regressed and experimental impedances; corresponding Bode plots of the first derivative of the exact, GP-regressed, and experimental impedances (c) and (d).
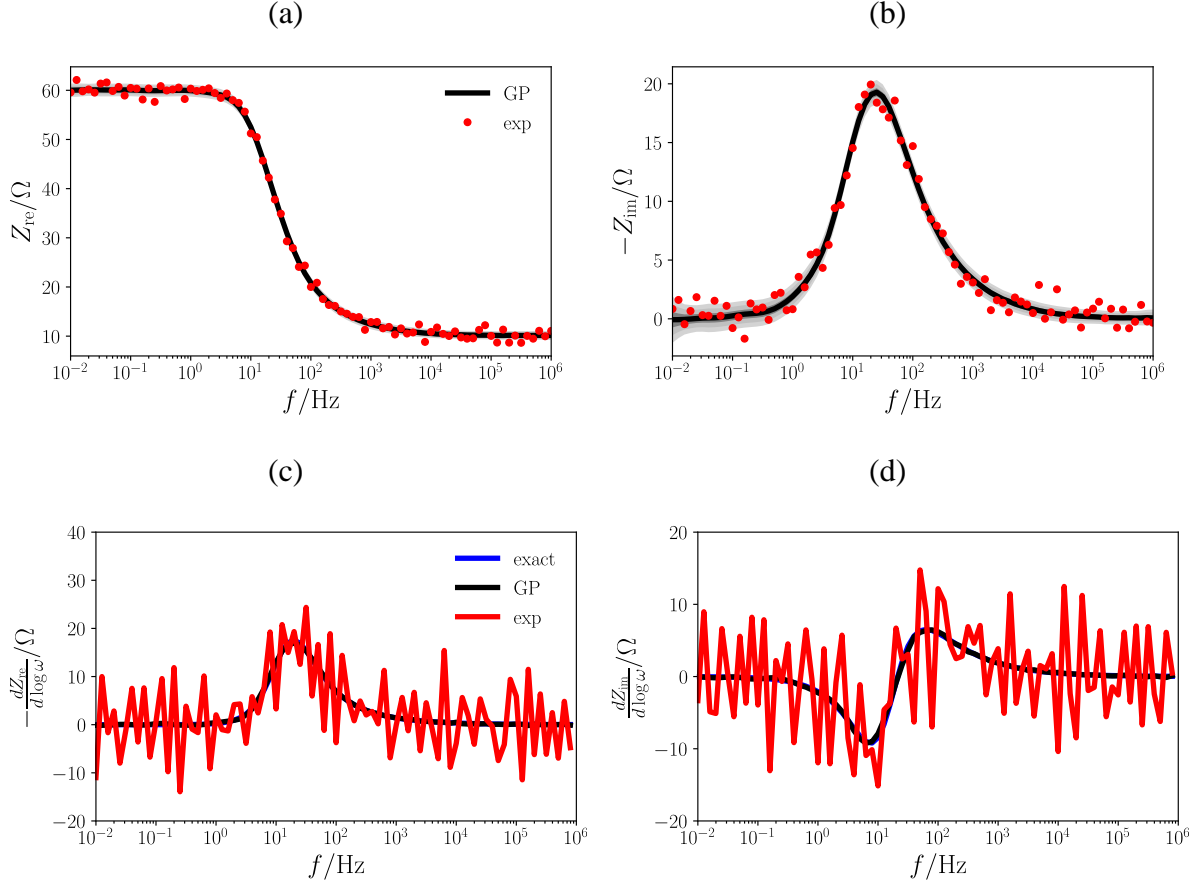
Figure S4. For $\sigma_n^{\mathrm{exp}} = 1.0\ \Omega$ and the fractal model, Bode plot of the (a) real and (b) imaginary part of the GP-regressed and experimental impedances; corresponding Bode plots of the first derivative of the exact, GP-regressed, and experimental impedances (c) and (d).

(a)

(b)

(c)

(d)

Figure S5. For $\sigma_n^{\text{exp}} = 1.0\ \Omega$ and the PWC model, Bode plot of the (a) real and (b) imaginary part of the GP-regressed and experimental impedances; corresponding Bode plots of the first derivative of the exact, GP-regressed, and experimental impedances (c) and (d).
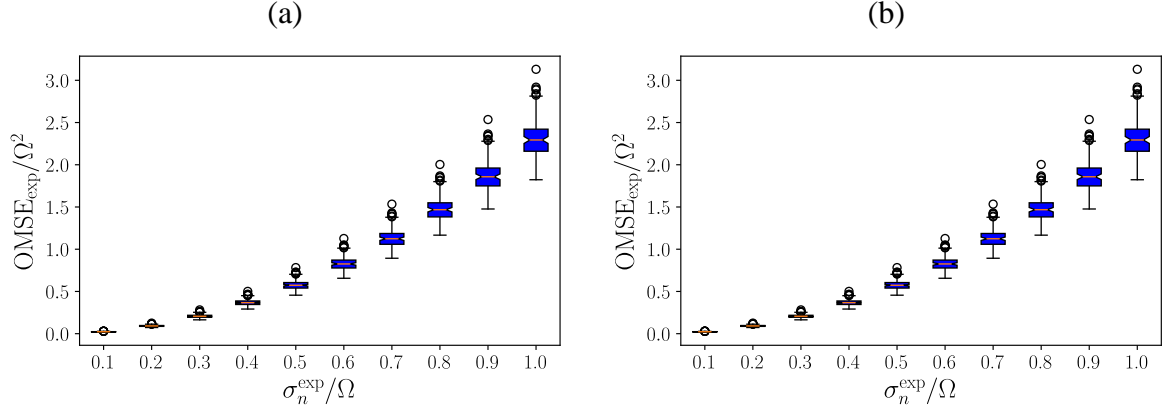
Figure S6. Boxplots of OMSE_{exp} as functions of σ_n^{exp} for the (a) 1×ZARC and (b) separated 2×ZARC models.
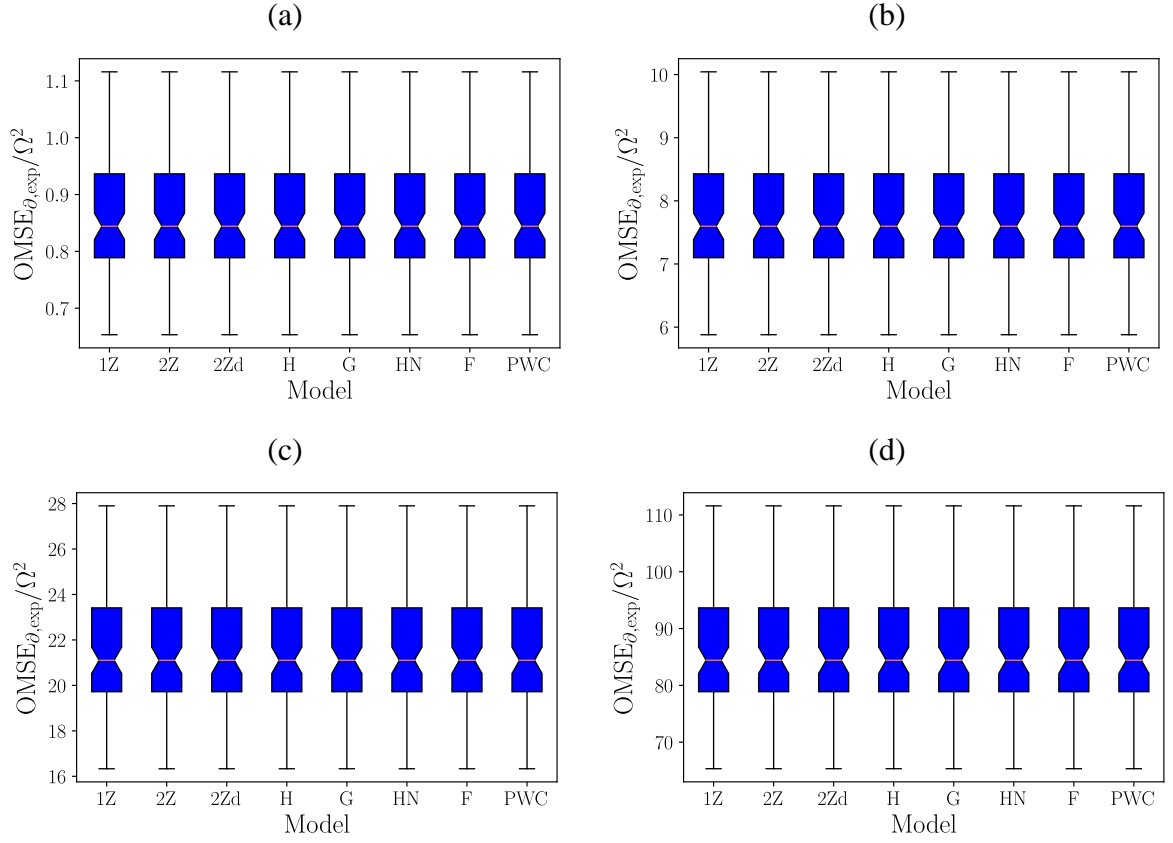

Figure S7. Boxplots of OMSE_{exp} for the 1×ZARC (1Z), separated 2×ZARC (2Z), distant 2×ZARC (2Zd), hook (H), Gerischer (G), Havriliak-Negami (HN), fractal (F), and PWC models for $\sigma_n^{\text{exp}} =$ (a) 0.1, (b) 0.3, (c) 0.5, and (d) 1.0 Ω.
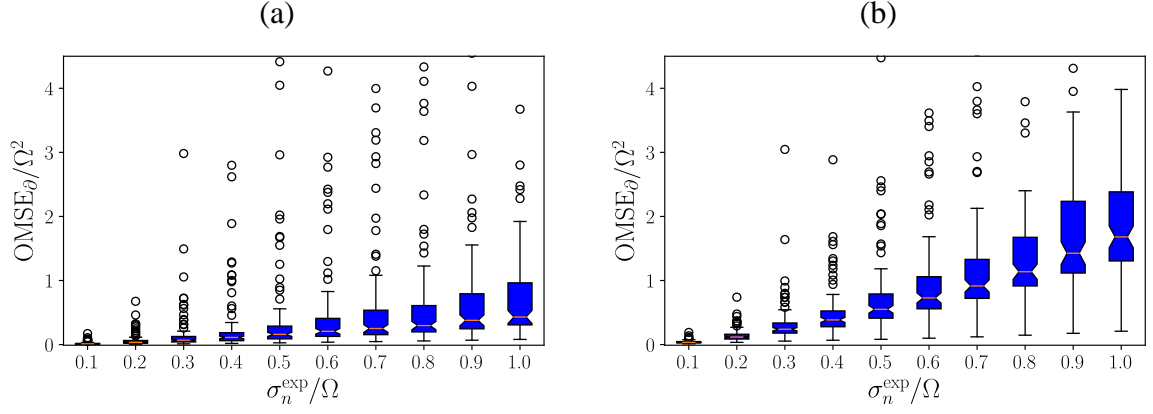
Figure S8. Boxplots of OMSE_∂ as functions of $\sigma_n^{\mathbf{exp}}$ for the (a) 1×ZARC and (b) separated 2×ZARC models.
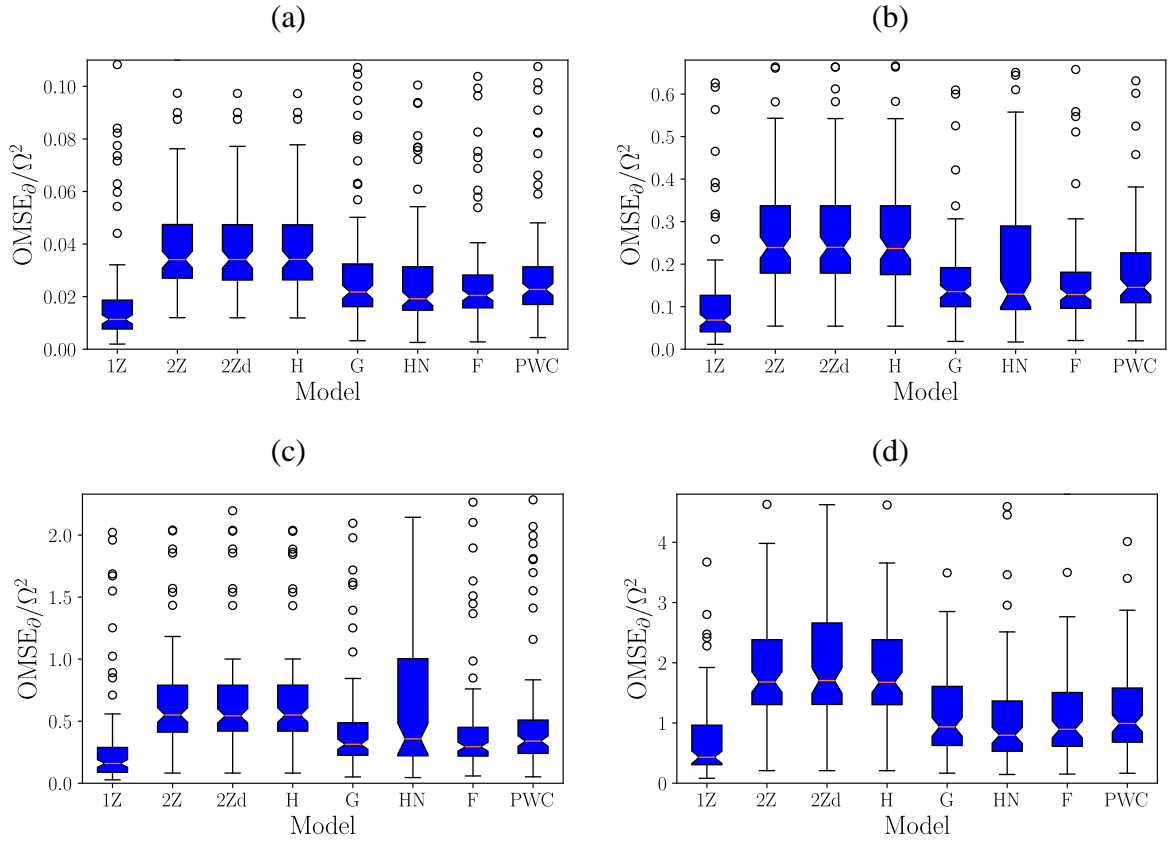
Figure S9. Boxplots of OMSE_∂ for the 1×ZARC (1Z), separated 2×ZARC (2Z), distant 2×ZARC (2Zd), hook (H), Gerischer (G), Havriliak-Negami (HN), fractal (F), and PWC models for $\sigma_n^{\mathbf{exp}} =$ (a) 0.1, (b) 0.3, (c) 0.5, and (d) 1.0 Ω.
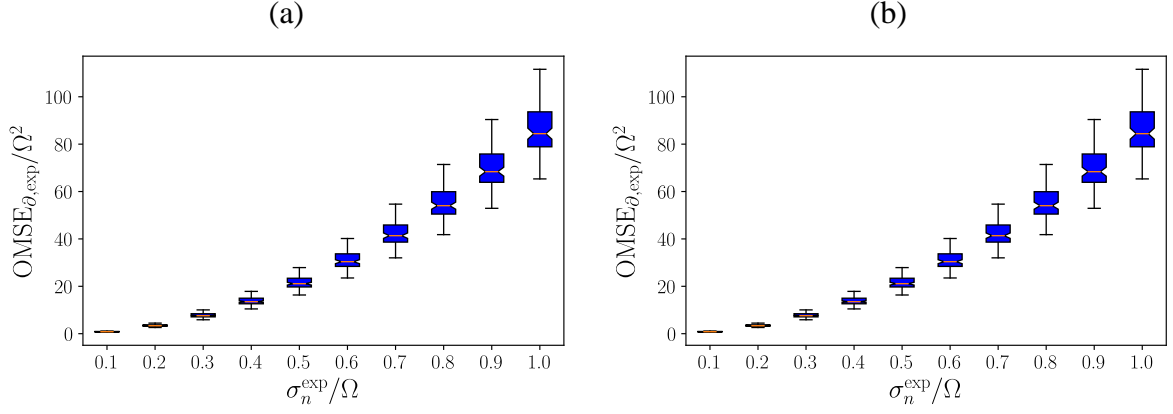
Figure S10. Boxplots of $\mathrm{OMSE}_{\partial,\mathrm{exp}}$ as functions of σ_n^{exp} for the (a) 1×ZARC and (b) separated 2×ZARC models.

Figure S11. Boxplots of $\mathrm{OMSE}_{\partial,\mathrm{exp}}$ for the 1×ZARC (1Z), separated 2×ZARC (2Z), distant 2×ZARC (2Zd), hook (H), Gerischer (G), Havriliak-Negami (HN), fractal (F), and PWC models for $\sigma_n^{\mathrm{exp}} =$ (a) 0.1, (b) 0.3, (c) 0.5, and (d) 1.0 Ω.
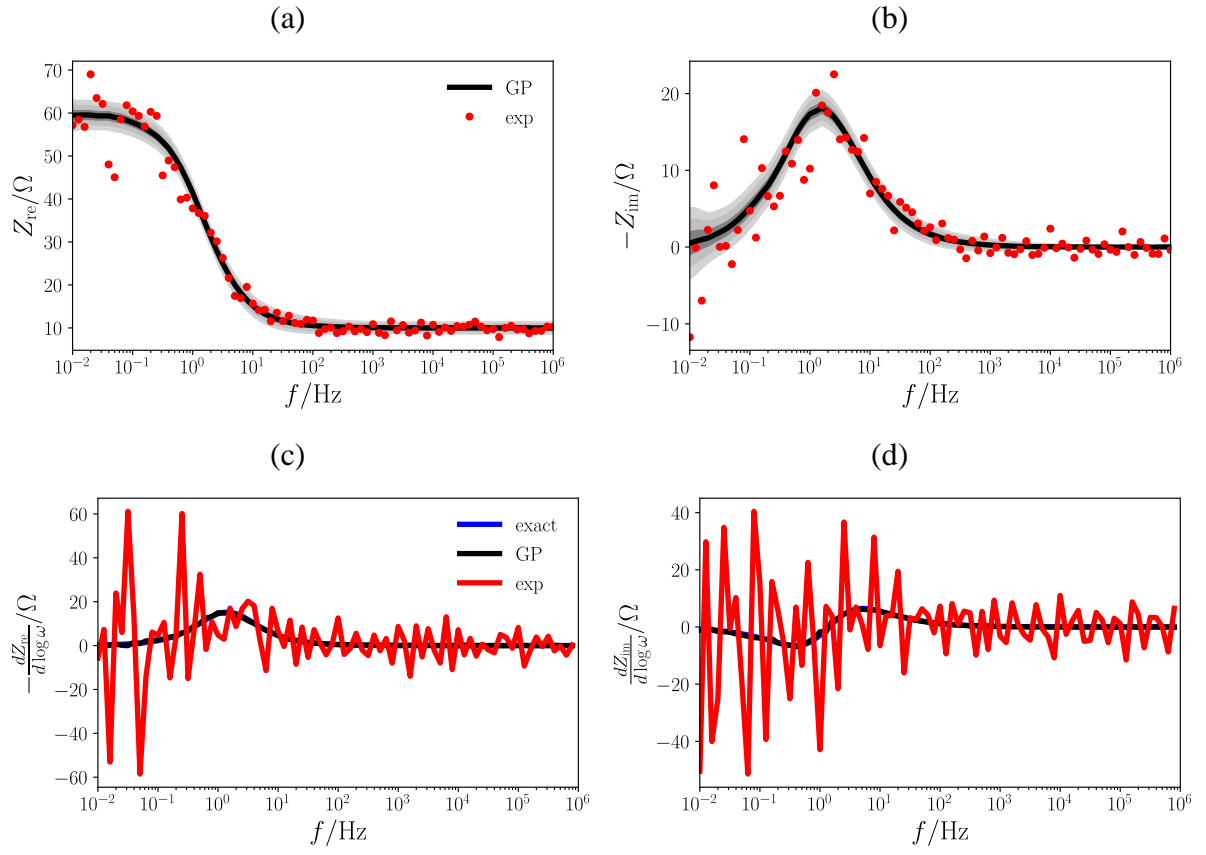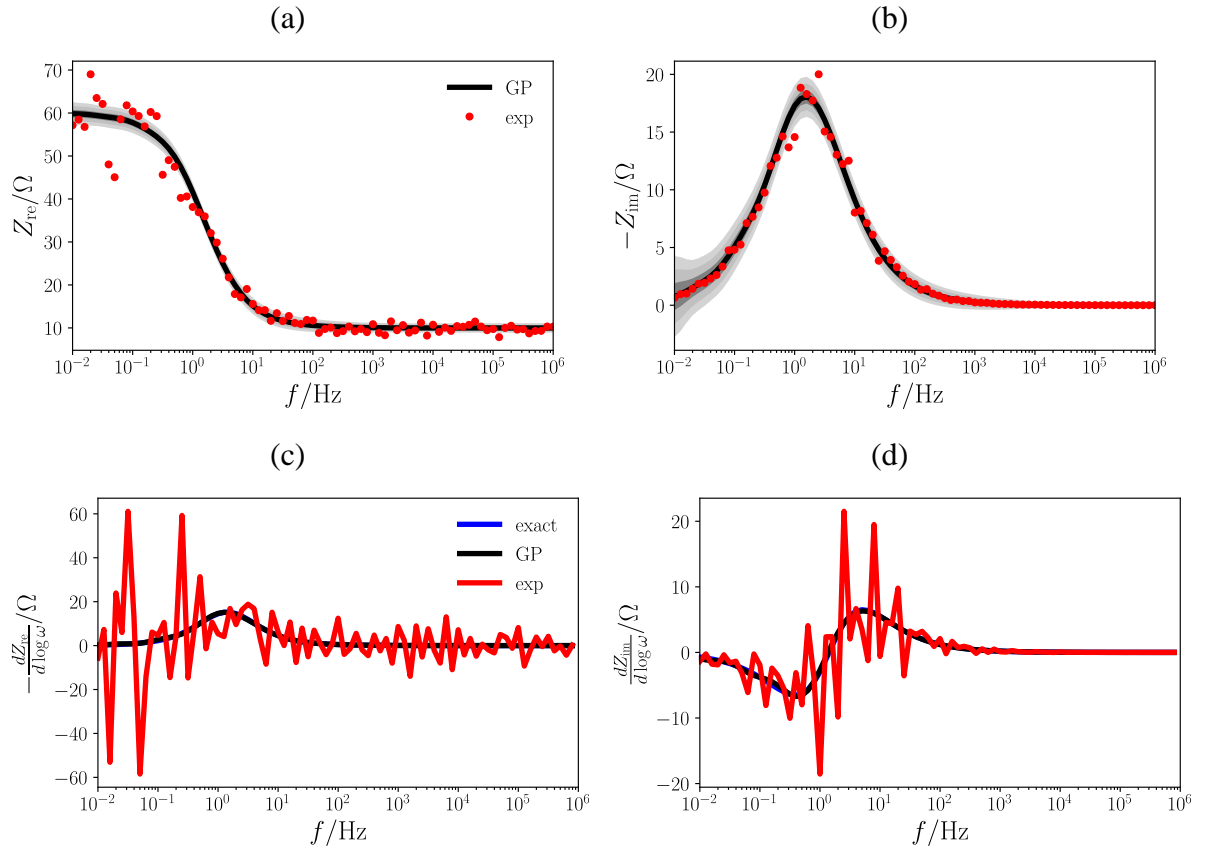
Figure S12. For the 1×ZARC model, Bode plot of the (a) real and (b) imaginary parts of the GP-regressed and experimental impedances for the noise model (24); corresponding Bode plots of the first derivative of the exact, GP-regressed, and experimental impedances (c) and (d).
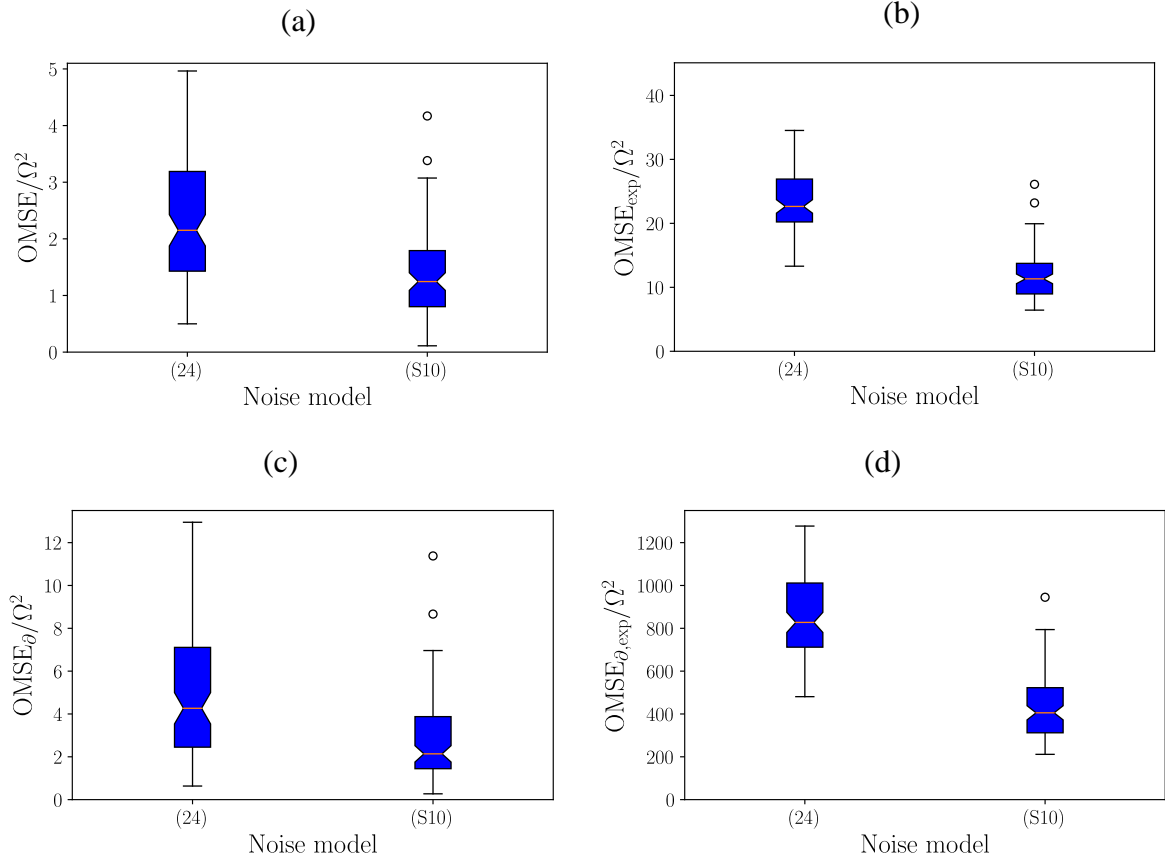
Figure S13. For the 1×ZARC model and the noise models (24) and (S10), boxplots of (a) OMSE, (b) OMSE_{exp}, (c) OMSE_{∂}, and (d) $\text{OMSE}_{\partial,\text{exp}}$.

Figure S14. For the 1×ZARC model, Bode plot of the (a) real and (b) imaginary parts of the GP-regressed and experimental impedances for the noise model (S10); corresponding Bode plots of the first derivative of the exact, GP-regressed, and experimental impedances (c) and (d).
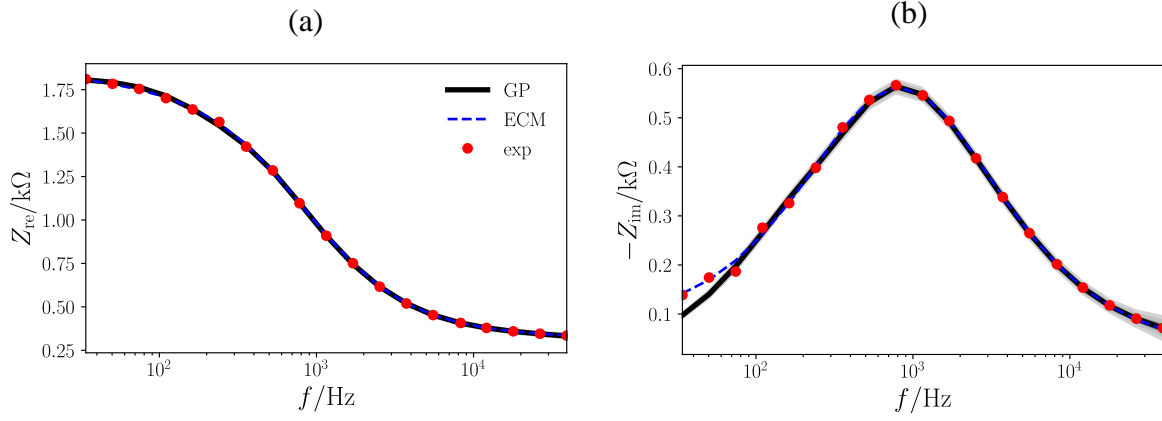
Figure S15. For the Li-metal battery, Bode plot of the (a) real and (b) imaginary parts of the GP-regressed, ECM, and experimental impedances.
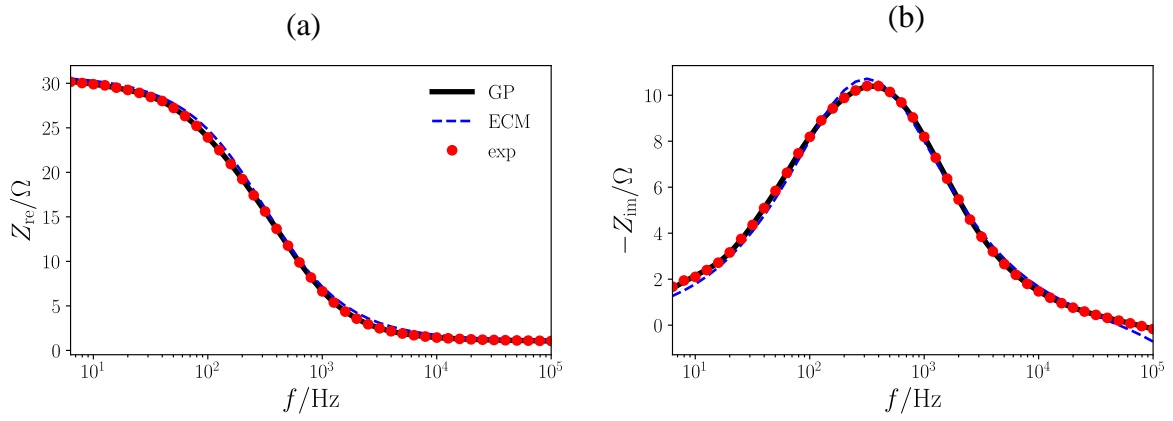

Figure S16. For the double-layer supercapacitor, Bode plot of the (a) real and (b) imaginary parts of the GP-regressed, ECM, and experimental impedances.
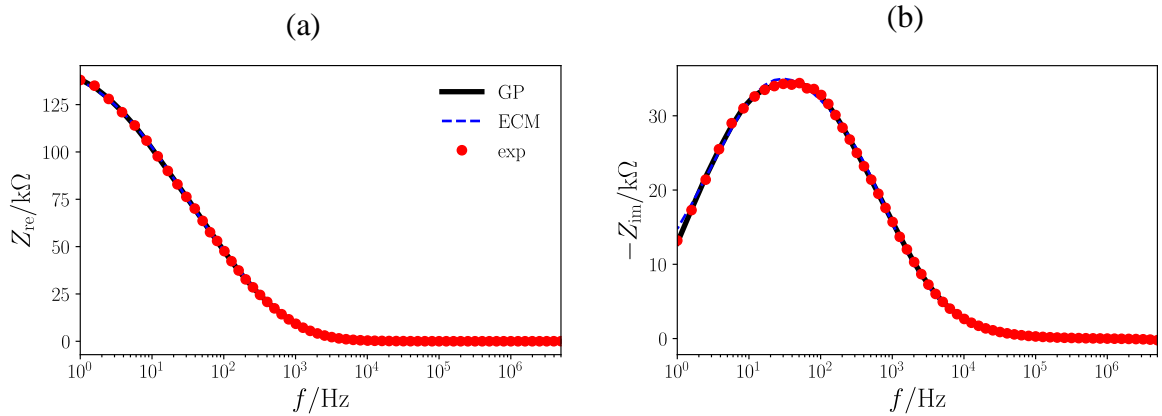

Figure S17. For the perovskite solar cell, Bode plot of the (a) real and (b) imaginary parts of the GP-regressed, ECM, and experimental impedances.
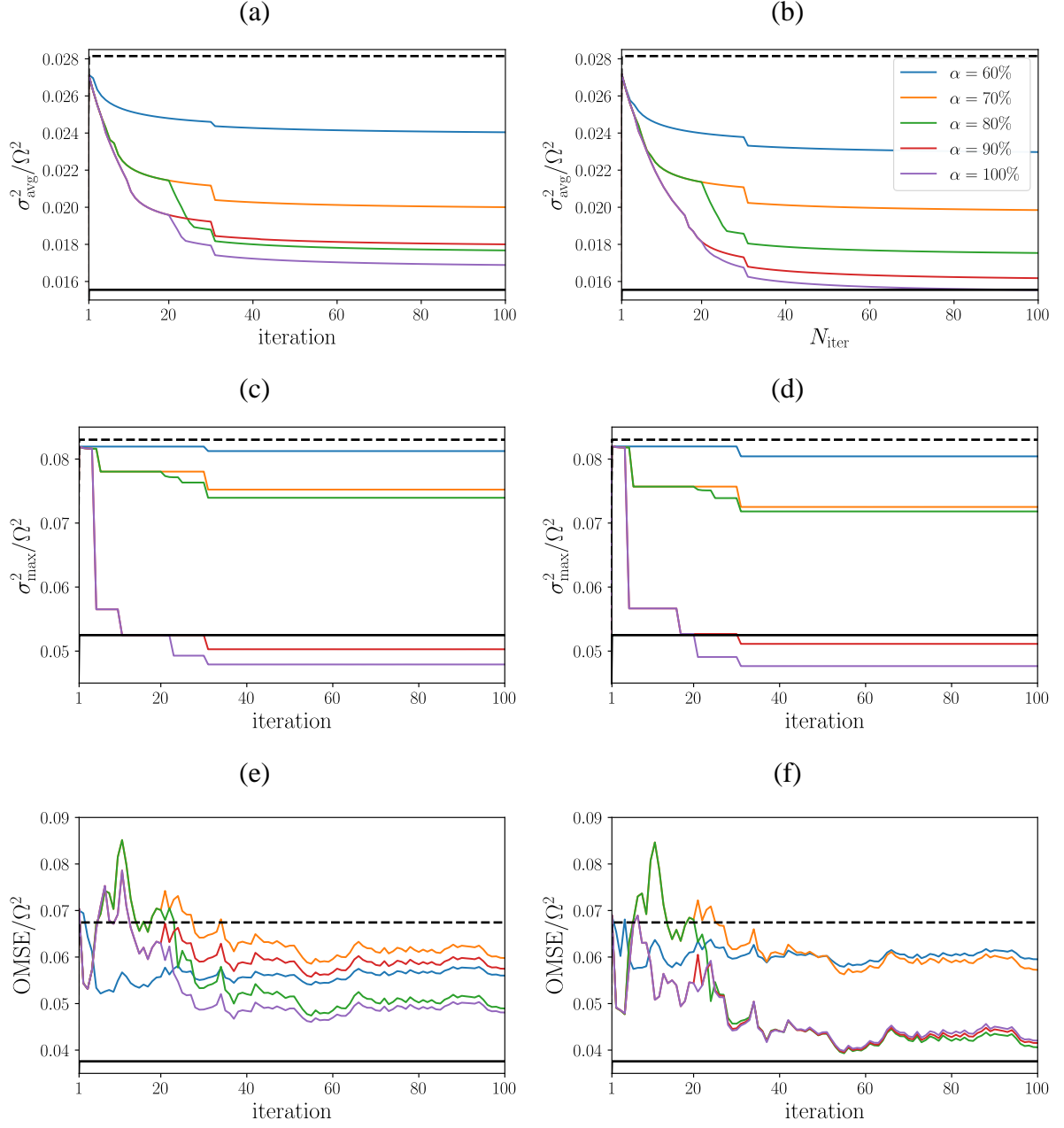
Figure S18. (a) σ_{avg}^2, (c) σ_{max}^2 , and (e) OMSE as functions of the iteration number for the D-criterion ALM; corresponding values for the D-criterion ALC are shown in panels (b), (d), and (f). The dashed and solid lines on each plot refer to the 5 and 10 ppd cases, respectively.
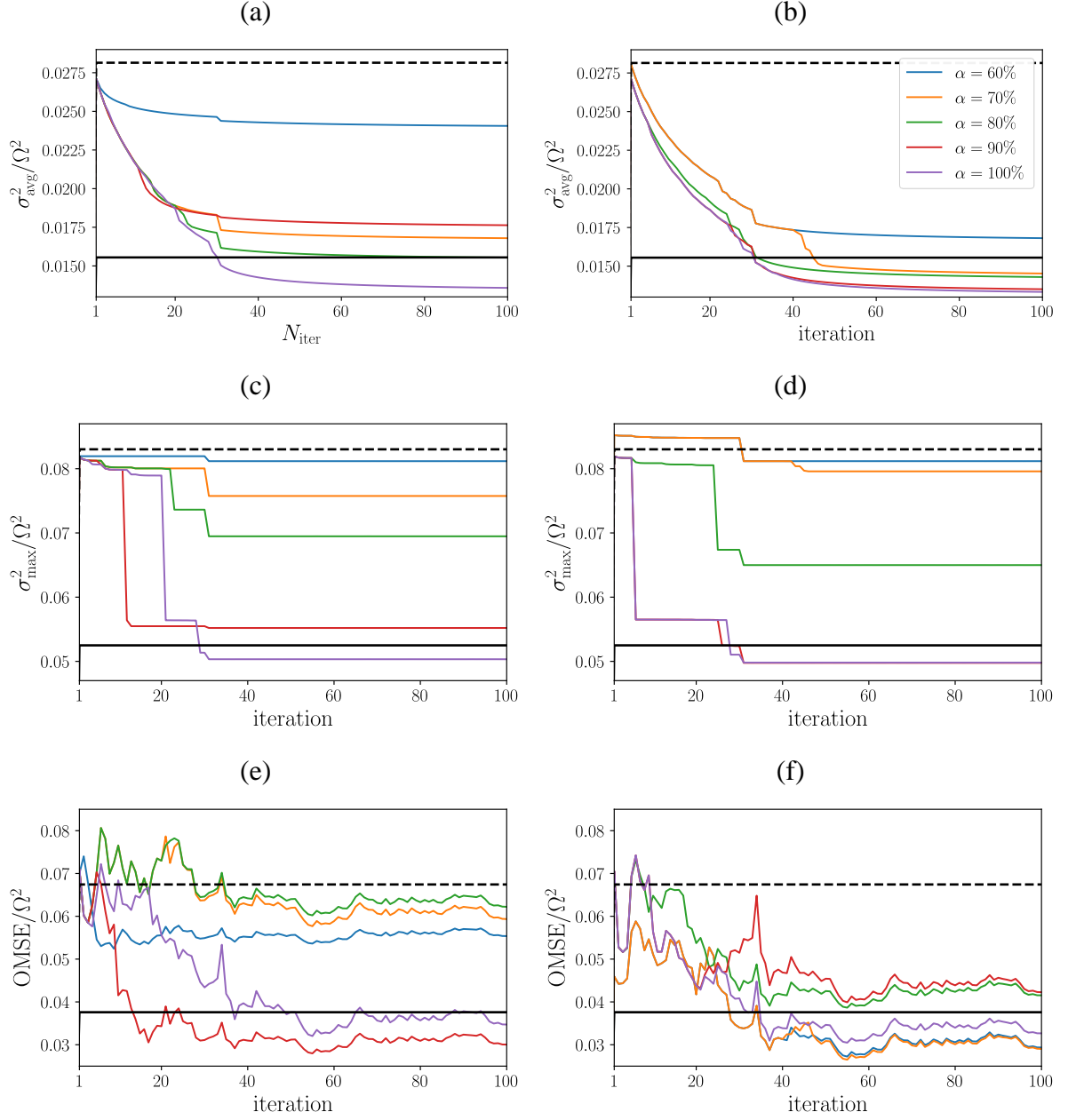
Figure S19. (a) σ^2_{avg}, (c) σ^2_{max}, and (e) OMSE as functions of the iteration number for the E-criterion ALM; corresponding values for the E-criterion ALC are shown in panels (b), (d), and (f). The dashed and solid lines on each plot refer to the 5 and 10 ppd cases, respectively.

References

[1] C.E. Rasmussen, C.K.I. Williams, Gaussian processes for machine learning, MIT Press, Cambridge, Mass, 2006.

[2] D.K. Duvenaud, Automatic model construction with Gaussian processes, Doctoral Dissertation, 157, University of Cambridge (2014).

[3] F. Ciucci, The Gaussian process Hilbert transform (GP-HT): Testing the consistency of electrochemical impedance spectroscopy data, J. Electrochem. Soc. 167 (2020) 126503.

[4] A. Gut, Probability: A graduate course, Springer 200 (2005) 5.

[5] P. Agarwal, O.D. Crisalle, M.E. Orazem, L.H. Garcia-Rubio, Application of measurement models to impedance spectroscopy: II . Determination of the stochastic contribution to the error structure, J. Electrochem. Soc. 142 (1995) 4149–4158.

[6] M.E. Orazem, P. Agarwal, C. Deslouis, B. Tribollet, Application of measurement models to electrohydrodynamic impedance spectroscopy, J. Electrochem. Soc. 143 (1996) 948–960.